THE UNIVERSITY *of* EDINBURGH

An observation of the charmless two-body decay $B \to \eta' K^*$ using data collected by the *BABAR* Experiment.

Alan Robertson

Doctor of Philosophy

The University of Edinburgh

2012

Abstract

A search for B decays to quasi two-body charmless final states involving a pseudoscalar η' meson recoiling against a K^* vector meson is described. This thesis primarily describes the analysis of two of the six possible decay channels, with the other four channels necessarily included as the subdecay modes are combined to give an overall branching fraction measurement. The method of analysis is a multivariate maximum likelihood fit for each subdecay channel. The likelihood curves for both modes are then combined, firstly with two other charged modes to yield an overall charged result, and finally the four charged modes are combined with two neutral modes to give an overall branching fraction and significance for the decay channel $B \to \eta' K^*$. All results use the full Run 1 to Run 4 datasets, comprising $210.5\,\text{fb}^{-1}$ of data, equivalent to 232 million $B\overline{B}$ pairs, gathered by the *BABAR* detector at Stanford Linear Accelerator Center in Menlo Park, California. The measured branching fractions and upper limits at 90% confidence limit (CL) are:

$$\begin{aligned} \mathcal{B}(B^+ \to \eta'_{\eta\pi\pi} K^{*+}_{K^+\pi^0}) &< 9.5 \times 10^{-6} \\ \mathcal{B}(B^+ \to \eta'_{\rho\gamma} K^{*+}_{K^+\pi^0}) &< 22 \times 10^{-6}. \end{aligned}$$

The four-mode combined fit determined the branching fraction for the decay $B^+ \to \eta' K^{*+}$:

$$\mathcal{B}(B^+ \to \eta' K^{*+}) < 7.9 \times 10^{-6}.$$

The six-mode combined fit determined the branching fraction for the decay $B \to \eta' K^*$:

$$\mathcal{B}(B \to \eta' K^*) = (4.1 \pm 1.0 \pm 0.5) \times 10^{-6}$$

at a significance of 5.6 standard deviations.

Declaration

This thesis has been composed by the author. The work contained within this thesis is either the author's own or that to which the author contributed significantly. This work has not been submitted for any other degree or professional qualification.

Signed:

.....

Alan Robertson

Acknowledgements

This thesis would not have been possible without the efforts of many individuals, organisations and institutions. The analysis was led by Dr Phil Clark (my long-suffering supervisor) and Dr Wolfgang Gradl. Their talents, plainly evident in bringing this analysis to publication, have clearly served them well in the period since this work was completed, and mean that I have pleasure in addressing them as Professor Phil Clark and Universitätsprofessor Wolfgang Gradl today. Sincere thanks, and many congratulations to you both on your recent appointments. Professor Steve Playfer has been a source of straight advice, in addition to patiently answering questions on accelerator physics and the detector. Much appreciated. I'd also like to thank Dr William Panduro Vazquez for taking the time to provide helpful comments on several chapters, and also for being an all-round good egg. Speaking of good eggs, Dr Debbie Bard, who also deserves a special mention for helping get this thesis kick-started. See what you've done now. I am also grateful to PPARC (R.I.P.) for funding this research.

There are too many individuals to credit individually with helping to ensure that life was never dull or short of friends in Edinburgh or beyond- you know who you are. That said, a special mention to the extremely talented Dr Eric Charles for ensuring a warm welcome awaited in California, and providing an exemplary display of how to get the most from life.

I'd like to end by thanking those closest to me. My parents have been very supportive of efforts to complete this project- thanks as always.

The thought that the *BABAR* experiment would be where I might meet my future wife didn't cross my mind when I started this caper. And then I met Katharine Schofield. Here's looking at you, kid.

Alan Robertson, 31st August 2012

Contents

List of Figures

List of Tables

Chapter 1

Theory

The existence and identification of symmetries in the natural world is useful to simplify models and further our understanding of the universe. Important symmetries in relation to the fundamental forces are charge conjugation(C), parity(P), and time reversal(T). Although the electromagnetic and strong force are found to be invariant under C, P and T transformations, the weak interaction was observed to violate parity conservation in nuclear β decays in 1957 [1], after Lee and Yang published a series of tests designed to highlight this behaviour [2]. In the following decade, combined CP symmetry was shown to be violated in weak decays in the kaon system [3]. This instigated much work, including the development of the Standard Model (SM). The SM accommodates CP violation as a consequence of a complex phase in the CKM matrix (see Section 1.1.3), but it does not predict the magnitude of the violation required to describe, for example, the matter/antimatter imbalance in the universe. Sakharov showed that CP violation is one of three conditions required in order to explain the universal matter/antimatter asymmetry [4].

The motivation to study CP violation in more detail is therefore clear, as it ought to further constrain the SM and give insights as to the direction in which any new physics beyond the SM should take if a comprehensive description of the universe is to be realised. B factories such as *BABAR* and Belle have measured and confirmed CP-violating decays in B mesons [5–8]. The attraction of examining CP violation at B factories such as PEP-II is

partly practical and partly scientific. Clean sources of B mesons are available, which is of obvious benefit in terms of data gathering. Additionally, the precision and accuracy with which theoretical parameters can be measured and constrained via the B meson system is superior to other systems. This is discussed in more detail below.

1.1 CP violation in the SM

1.1.1 Quark mixing in weak interactions

In contrast to strong decays, weak decays result in quark mixing. This was concluded by the observation that the Fermi coupling constant, G_F , accurately derived from the $d \to u$ transition in muon lifetime measurements, appeared to vary in other weak decay channels where the final state quark was different, e.g. the $s \to u$ in kaon decays. Cabibbo concluded that the relative strength of coupling was dependent on the mixture of d and s mass eigenstates which couple to the d' and s' weak eigenstates, and that the magnitude of the coupling could be described by a single parameter: the Cabibbo angle, θ [9]. Additional work formulated a system in which the mixing of the lightest four quarks could be described by the Cabibbo angle as follows [10]:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}. \tag{1.1}$$

This formalism allowed for the observed mixing between quarks of differing charge, whilst preserving the observed near zero mixing associated with neutral currents. However, this formalism could not account for CP violation, as the couplings described are dependent on the single, real Cabibbo angle and are unaffected by reversals in CP symmetry. CP violation could not be described within a four quark system.

1.1.2 The SM description

The SM describes CP violation as arising from a single phase in the quark mixing matrix [11]. The Lagrangian describing the weak interaction is as follows:

$$\mathcal{L}_{weak} = \frac{g}{\sqrt{2}} \left(G^{\mu}_{weak} W^{\pm}_{\mu} + G^{\mu\dagger}_{weak} W^{\mp}_{\mu} \right) \tag{1.2}$$

where the weak current operator G^{μ}_{weak} is equivalent to:

$$G^{\mu}_{weak} = \bar{U}'_L \gamma^{\mu} D'_L \tag{1.3}$$

where $\bar{U}_L$ and D'_L are the left-handed components of the up- and down-type quark fields. γ^{μ} are the gamma matrices for $\mu = 0, 1, 2, 3$. G^{μ}_{weak} has nine parameters; three angles may be defined, in an analogous manner to the Cabibbo angle in the four quark system; the remaining six are interdependent phases. Five of these phases, describing transitions between the six quarks, can be eliminated by transformations, leaving one:

$$G^{\mu}_{weak} \Rightarrow P_u G^{\mu}_{weak} P^*_d \tag{1.4}$$

where P_u and P_d are diagonal matrices which redefine the phase of the quark mass eigenstates. The remaining complex phase allows CP violation, and is known as the Cabibbo-Kobayashi-Maskawa (CKM) phase [12]. The transformed matrix is the CKM matrix.

1.1.3 The CKM matrix

The CKM matrix approach with three generations was developed to accommodate CP violation in the SM, while the 2×2 Cabibbo-angle based mixing matrix could not account for this. The transformation in Equation 1.4 allows CP violation via the complex phase, coupling the mass eigenstates of quarks to weak eigenstates, which does vary under CP symmetry [12]. The CKM matrix is a 3×3 complex, unitary matrix shown in Equation 1.5:

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{1.5}$$

This representation shows each element in terms of the related quarks. The Particle Data Group (PDG) [13] parameterisation of the CKM matrix is shown in Equation 1.6:

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}c_{23}s_{13}e^{-i\delta} & c_{12}c_{23} - s_{12}s_{23}e^{-i\delta} & s_{23}c_{13} \\ s_{12}c_{23} - c_{12}c_{23}s_{13}e^{-i\delta} & -c_{12}c_{23} - s_{12}s_{23}e^{-i\delta} & c_{23}c_{13} \end{pmatrix} \tag{1.6}$$

where c_{ij}, s_{ij} are the cosine and sine of the free parameters, θ_{ij}, and $e^{-i\delta}$ is the phase responsible for CP violation. The matrix is parameterised by three angles and the (non-zero) complex phase. An alternative parameterisation of the CKM matrix is the Wolfenstein parameterisation, which makes use of the fact that c_{23} and c_{13} are ≈ 1. Also, defining $\lambda = s_{12} \approx 0.22$ we arrive at the parameterisation shown in Equation 1.7:

$$V = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} \tag{1.7}$$

where A, ρ and η are real numbers of order 1. This parameterisation is useful as it uses the quantities ρ and η which are accessible experimentally. This can be seen by considering one of the unitarity constraints on the matrix, known as the unitarity triangle.

1.1.4 The $B^0 - \overline{B^0}$ unitarity triangle

We can use the CKM matrix as shown in Equation 1.5 and the unitarity condition $V^\dagger V = \mathcal{I}$ to write down nine equations which must hold. One of these is of particular interest to the *BABAR* experiment as it includes $V_{tb}^\dagger V_{td}$ which describes the $b \to d$ transition responsible for $B^0 - \overline{B}^0$ mixing. It is shown in Equation 1.8:

$$V_{ub}^* V_{ud} + V_{cb}^* V_{cd} + V_{tb}^* V_{td} = 0. \tag{1.8}$$

This can be graphically represented as a triangle in the complex plane since the three terms sum to zero. The fact the triangle has a non-zero area allows CP violation to occur. It is conventional to normalise the three terms to $|V_{cb}^* V_{cd}|$, with an appropriate choice of phase to make the $V_{cb}^* V_{cd}$ term real (in graphical terms, rotating the diagram onto the real axes so the base of the triangle lies on the real axis). The unitarity triangle is shown in Figure 1.1.

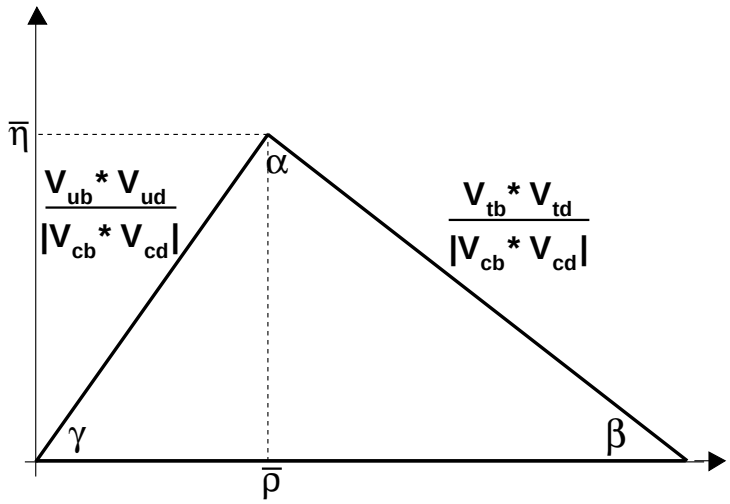

Figure 1.1: The unitarity triangle. In this representation, the location of the only free vertex is given in terms of the Wolfenstein parameters at $(\overline{\rho}, \overline{\eta})$, with $\overline{x} = x(1 - \lambda^2/2), x = (\eta, \rho)$, a small correction which enhances the accuracy of the parameters. The other vertices are fixed at $(0, 0)$ and $(0, 1)$.

The angles of the unitarity triangle are calculated as follows:

$$\alpha = -arg\left(\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right) \tag{1.9}$$

$$\beta = -arg\left(\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) \tag{1.10}$$

$$\gamma = -arg\left(\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \tag{1.11}$$

The aim of the *BABAR* experiment is to measure or constrain as many sides and angles of the unitarity triangle as possible. This enables the SM theory of CP violation to be tested. The current status of the constraints on the angles and sides is shown in Figure 1.2.

It is possible to draw six unitarity triangles from Equation 1.5, each of which can be used to examine CP violation. Only two of the equations yield triangles which have roughly equal side lengths, and these relate to the B meson system. The relative side lengths will be roughly determined by the exponent of λ in each of the terms in Equation 1.8. Since λ is small, the significance of terms, and hence side lengths in the graphical representation, will diminish rapidly relative to others with increasing powers of λ. The other graphical representations, relating to the D meson and kaon systems, are more uneven graphically which makes it more difficult to measure the angles with precision since they can be very small. *BABAR* exploits the triangle which relates to the B_d which also has the advantage that a clean source of mesons is available via the $\Upsilon(4S)$ resonance. The other B meson triangle, corresponding to the B_s system, also has a convenient graphical parametrisation, but the higher mass of this meson is above the centre of mass energy production threshold for the *BABAR* experiment.

1.2 Methods of CP violation

There are three means by which CP symmetry may be violated:

- by decay, which can occur for both charged and neutral particles,
- via mixing, restricted to neutral particles due to charge conservation,
- via interference, also limited to neutral particles.

Each is described in more detail below.

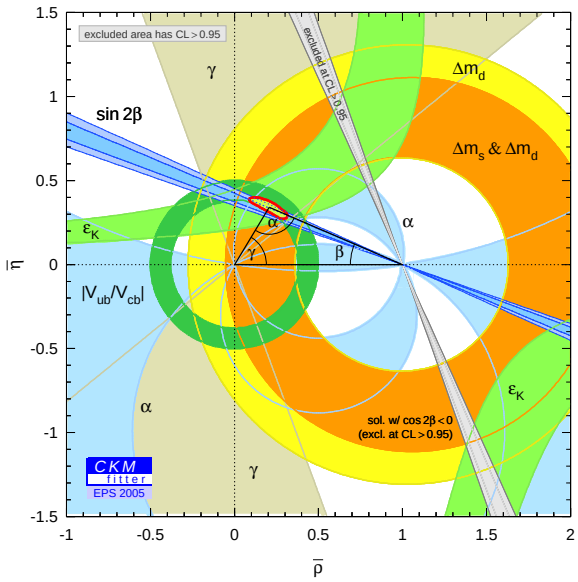

Figure 1.2: CKM constraints shown in the plane of the Wolfenstein parameters $(\overline{\rho}, \overline{\eta})$.

1.2.1 CP violation in decay

This is often referred to as "direct" CP violation, and can occur in both charged and neutral decays. CP violation in decay occurs when the decay amplitude and its CP conjugate differ. This is a direct result of a complex contribution to the decay amplitudes, as this will contribute with opposite sign to each of the conjugate processes.

Consider the neutral decays of R^0 and its antiparticle $\overline{R^0}$ into a final state, f. The decays are described by two decay amplitudes:

$$A_f \equiv \langle f|H|R^0\rangle = \Sigma_i A_i e^{i(\delta_i+\phi_i)} \tag{1.12}$$

$$\overline{A_f} \equiv \langle f|H|\overline{R^0}\rangle = \Sigma_i A_i e^{i(\delta_i-\phi_i)} \tag{1.13}$$

where $e^{i\delta_i}$ and $e^{i\phi_i}$ are the strong and weak phase terms respectively. $|A_f|$ and $|\overline{A_f}|$ will be equal (CP symmetry conserved) when the weak phases, ϕ_i , are the same. Otherwise:

$$|A_f| \neq |\overline{A_f}| \tag{1.14}$$

and CP symmetry is violated. It can be seen from Equations 1.12 and 1.13 that CP violation of this type can only occur when at least two terms that have different weak phases also have different strong phases:

$$|A^2| - |\overline{A}|^2 \equiv -2\sum_{ij} A_i A_j \sin(\phi_i - \phi_j)\sin(\delta_i - \delta_j) \tag{1.15}$$

since trivially if either weak or strong phases are equivalent the RHS is zero. In charged decays, relevant for this analysis, the asymmetry is defined as:

$$a_f = \frac{\Gamma(B^+ \to f) - \Gamma(B^- \to \overline{f})}{\Gamma(B^+ \to f) + \Gamma(B^- \to \overline{f})} = \frac{1 - |\overline{A}/A|^2}{1 + |\overline{A}/A|^2}. \tag{1.16}$$

1.2.2 CP violation in mixing

CP violation in mixing involves neutral particles only where they oscillate to their antiparticle. The violation arises due to the particles' mass eigenstates being different from their CP eigenstates. This can be examined by considering the following system of a linear combination of neutral B meson eigenstates:

$$a|B^0\rangle + b|\overline{B^0}\rangle. \tag{1.17}$$

The time evolution of the system is governed by the time-dependent Schrodinger equation:

$$i\frac{d}{dt}\begin{pmatrix} a \\ b \end{pmatrix} = \mathcal{H}\begin{pmatrix} a \\ b \end{pmatrix} \equiv (M - \frac{i}{2}\Gamma)\begin{pmatrix} a \\ b \end{pmatrix} \tag{1.18}$$

where M and Γ are 2×2 Hermitian matrices. CPT invariance imposes the conditions $\Gamma_{11} = \Gamma_{22}$. The off diagonal elements M_{12} and Γ_{12} are responsible for CP violation in the transitions between B^0 and $\overline{B^0}$.

The mass eigenstates of the system, denoted B_L and B_H are given by:

$$|B_L\rangle = p|B^0\rangle + q|\overline{B^0}\rangle \tag{1.19}$$

$$|B_H\rangle = p|B^0\rangle - q|\overline{B^0}\rangle \tag{1.20}$$

where p and q are complex coefficients which obey the normalisation condition:

$$|p|^2 + |q|^2 = 1. \tag{1.21}$$

Solving equation 1.18 allows the ratio q/p to be expressed:

$$\frac{q}{p} = \frac{\Delta m_B - \frac{i}{2}\Delta\Gamma_B}{2(M_{12} - \frac{i}{2}\Gamma_{12})} = \frac{2(M_{12}^* - \frac{i}{2}\Gamma_{12}^*)}{\Delta m_B - \frac{i}{2}\Delta\Gamma_B}. \tag{1.22}$$

To examine the conditions where CP symmetry will be violated, we look at the following quantity:

$$(q/p)^2 = \frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}. \tag{1.23}$$

When the mass eigenstates are also CP eigenstates, the relative phase between them disappears and CP symmetry holds. CP in mixing occurs when $|q/p| \neq 1$. This can be observed in the neutral B system in semileptonic decays with the asymmetry given by:

$$A_{CP}(t) \equiv \frac{\Gamma(\overline{B^0}(t) \to l^+\nu X) - \Gamma(B^0(t) \to l^-\overline{\nu} X)}{\Gamma(\overline{B^0}(t) \to l^+\nu X) + \Gamma(B^0(t) \to l^-\overline{\nu} X)}, \tag{1.24}$$

which can be written in terms of q and p as follows:

$$a_{sl} = \frac{1 - |q/p|^4}{1 + |q/p|^4}. \tag{1.25}$$

The effects of CP violation in mixing are expected to be small ($\mathcal{O}(10^{-2})$). Also, there are large uncertainties in calculating Γ_{12} which make it difficult to relate the CP violation observed to CKM parameters.

1.2.3 CP violation in interference between decay with and without mixing

This occurs in neutral decays to CP eigenstates which are accessible to both the B^0 and $\overline{B}^0$. In direct CP violation and CP violation in mixing it was shown that CP violation

occurred when $|A_f| \neq |\overline{A_f}|$ in decay and $|q/p| \neq 1$ in mixing. However, it is still possible for CP violation to occur even when $|A_f| = |\overline{A_f}|$ in decay and $|q/p| = 1$. We define a phase-independent quantity, λ:

$$\lambda_{fCP} \equiv \frac{q}{p}\frac{\overline{A}_{fCP}}{A_{fCP}} = \eta_{fCP}\frac{q}{p}\frac{\overline{A}_{\bar{f}CP}}{A_{fCP}} \tag{1.26}$$

with

$$\overline{A}_{fCP} = \eta_{fCP}\overline{A}_{\bar{f}CP} \tag{1.27}$$

where η_{fCP} is the CP eigenvalue of the state f_{CP} and has possible values of ± 1. CP violation occurs when $|\lambda| \neq 1$, but it can also happen if the imaginary part of λ is non-zero. This is CP violation in interference between decays with and without mixing. We can define the CP asymmetry of a time-dependent system:

$$A_{CP}(t) \equiv \frac{\Gamma(B^0(t) \to f) - \Gamma(\overline{B^0}(t) \to f)}{\Gamma(B^0(t) \to f) + \Gamma(\overline{B^0}(t) \to f)}. \tag{1.28}$$

1.3 Rare hadronic B decays

Rare hadronic B decays are of interest as they allow precision measurements of some of the CKM matrix elements, which allows the SM to be constrained. If the constraints on the SM become inconsistent, new physics beyond the SM (e.g. supersymmetry) is implied. The importance of rare hadronic B decays comes from the role of Flavour Changing Neutral Currents (penguin diagrams are an example of a FCNC) in these decays, which are forbidden at tree level in the SM. Interference produced by penguin contributions may result in a CP-violating asymmetry greater than the SM predictions, another indicator of new physics. This thesis is concerned with one such rare decay, discussed below.

1.3.1 The decay $B^+ \to \eta' K^{*+}$

This analysis is concerned with two of the four sub-decays of the final state $\eta' K^{*+}$, namely those within which the K^{*+} decays to a charged kaon and a π^0. The η' decays to either $\eta\pi\pi$ or $\rho\gamma$. The primary mesons of interest to this analysis are the B^+, η' and K^{*+}. The quark content of each is shown in Table 1.1.

This decay is of interest partly due to the presence of the η', which is closely related to the η meson as they are orthogonal combinations of two flavour-SU(3) states, η_1 and η_8, which mix. The singlet and octet are defined as follows:

$$\eta_1 = \frac{1}{\sqrt{3}}(u\overline{u} + d\overline{d} + s\overline{s}) \tag{1.29}$$

$$\eta_8 = \frac{1}{\sqrt{6}}(-u\overline{u} - d\overline{d} + 2s\overline{s}). \tag{1.30}$$

The mixing angle has been determined to be around twenty degrees [14]. However, a straightforward approach to $\eta - \eta'$ mixing does not explain the phenomenon whereby some final states are suppressed (e.g. $\eta' K^*$) and others enhanced (e.g. $\eta' K$). Lipkin argues that the modelling of the η and η' as mixtures of ground state $q\overline{q}$ systems is overly simple, and that radially excited $q\overline{q}$ systems must be considered if a satisfactory explanation for the observed behaviour is to be formulated [15]. Radially excited $q\overline{q}$ systems affect the behaviour of the η and η' wavefunctions at high momentum and allow diagrams requiring high momentum transfers to contribute in a manner they could not in the ground-state model.

The current theoretical predictions for this decay and the experimental status before this analysis are shown in Table 1.2, and a brief discussion of the theoretical approaches follows. There is predicted to be very small CP asymmetry in this decay channel. The motivation for studying this decay is clear, as there are theoretical predictions, but few experimental measurements. The *BABAR* experiment's contribution was achieved with $82\,\text{fb}^{-1}$ [16]. The data-set has since been increased to $210\,\text{fb}^{-1}$, and this provided a clear opportunity to further constrain the upper limit or to make a measurement and test the theoretical approaches. For

Meson	Quark content
B^+	$u\bar{b}$
η'	$\eta_1 \cos\theta + \eta_8 \sin\theta$
η	$\eta_8 \cos\theta$ - $\eta_1 \sin\theta$
K^{*+}	$u\bar{s}$
π^0	$\frac{1}{\sqrt{2}}(-u\bar{u} + d\bar{d})$

Table 1.1: Quark content of B^+, η', η, K^{*+} and π^0 mesons. The η' and η mesons are superpositions of SU(3) octet (η_8) and singlet (η_1) states.

Decay mode	Theoretical Predictions ($\times 10^{-6}$)		Experimental status ($\times 10^{-6}$)		
	SU(3) flavour [18]	QCD fact. [19]	HFAG(7/05) [17]	*BABAR* [16]	Belle [20]
$B^+ \to \eta' K^{*+}$	$2.8^{+1.2}_{-0.3}$	$5.1^{+10.31}_{-\ 5.94}$	< 14	< 14	< 90

Table 1.2: Theoretical predictions and experimental status of the decay $B^+ \to \eta' K^{*+}$ prior to the analysis presented here. Where only upper limits exist in the experimental results, the HFAG take the lowest of these as the official value.

decay channels where upper limits are the only experimental contributions available, only the best upper limit is quoted by the Heavy Flavor Averaging Group (HFAG) [17].

1.3.2 Contributing diagrams

The main contributing diagrams are shown in Figure 1.3. The $b \to s$ gluonic penguin diagrams (top row) may be expected to dominate. This is a result of the CKM supression of the external tree diagram due to the size of V_{ub}, equal to λ^3 in the Wolfenstein parameterisation. The gluonic penguin in which the η' is formed purely from hadronisation is often considered to be suppressed due to the "OZI rule" [21], and thus neglected in calculations of the decay amplitude by many theoretical treatments. However this diagram is of key interest to the explanation of the $\eta - \eta'$ puzzle, as the non-standard treatment of the $\eta - \eta'$ mixing angle de-

scribed in [15] demonstrates that the OZI-suppressed diagrams can contribute significantly to the observed enhancement seen in some final states, since the η and η' wavefunctions' high-momentum behaviour is altered.

Lastly, there are also internal tree diagrams (middle row) which are both CKM and colour suppressed (due to the colour singlets which need to be formed from quarks of independent colour structure), and a variety of possible singlet diagrams.

Together the diagrams mentioned form the basis for the SU(3) symmetry theoretical predictions. The singlet contribution for the η', defined as the amplitude associated with producing the quark-antiquark or gluon pair which then hadronises in the coherent flavour state of the η' has been the subject of much work [22]; it requires special treatment in QCD factorisation since the standard formulae do not hold.

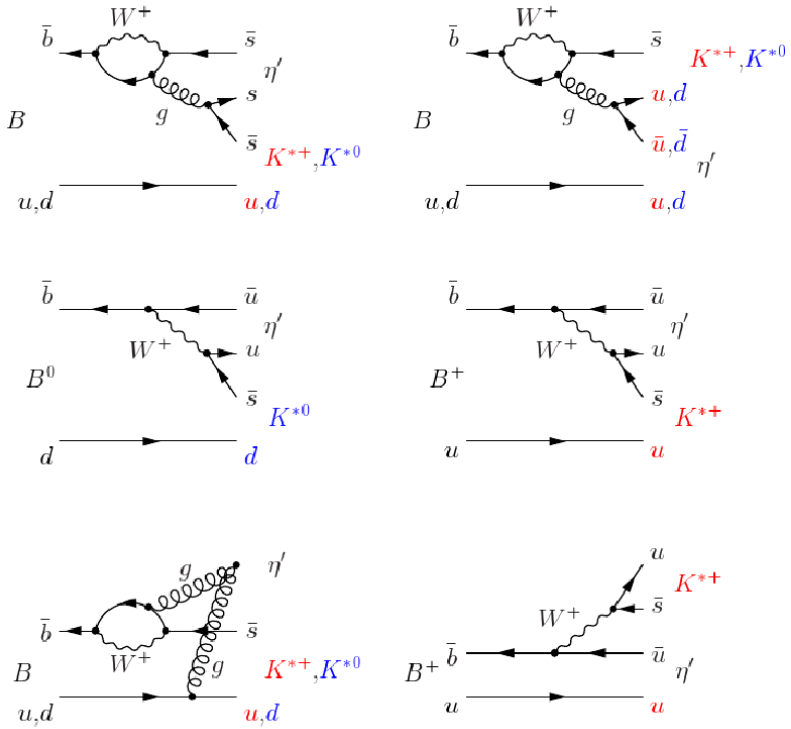

Figure 1.3: Feynman diagrams for the decays $B^+ \to \eta' K^{*+}$and $B^+ \to \eta' K^{*0}$.

1.4 Theoretical approaches

1.4.1 Flavour SU(3) symmetry

The decay amplitudes calculated by theory using the SU(3) symmetry approach are calculated from quark-level diagrams [18, 23–25], the most significant of which are shown above. This allows relations between the various contributing channels to be quantified. A fundamental assumption is that the light quark triplets are indistinguishable during the action by the strong force. The weak phase seen in Equation 1.12 is a fitted parameter, relevant to both branching fractions and CP asymmetries.

The approach is attractive due to the intuitive nature of the contributions to the decay amplitude from each diagram, however the drawbacks are that SU(3) symmetry is not exact since the quarks in the triplet have differing masses. Also the matrix elements of the Hamiltonian are not calculated explicitly as in the effective Hamiltonian approach. Instead a series of relations between different decays with similar components is established which constrains the decay amplitude of interest. This means that to make quantitative predictions, experimental input is required in the fits to determine the absolute decay amplitudes for the modes. The relative relations provided by SU(3) flavour can be invoked to make predictions for other modes.

A major advantage of the SU(3) approach to decay channels involving η' is its more sophisticated treatment of $\eta - \eta'$ mixing, mentioned above. The contributions to the amplitude of the $B^+ \to \eta' K^{*+}$ decay come from the diagrams shown in Figure 1.3 and are as follows [18]:

$$A(B^+ \to \eta' K^{*+}) = \frac{1}{\sqrt{6}}(p'_p + 2p'_v + t'_p + c'_v + 4s'_v) \tag{1.31}$$

where p'_p is the penguin diagram where the spectator quark forms the pseudoscalar (η') meson's amplitude (top right diagram in Figure 1.3), p'_v is where the spectator quark forms the vector (K^*) meson (top left), t'_p is the external tree-level contribution (bottom right) and c'_v and s'_v are the colour suppressed and singlet contributions. The penguin contributions are the most significant, but since it is assumed that $p'_p = -p'_v$, in this decay the QCD penguin

contributions will destructively interfere and hence the predicted branching fraction is small. This is in contrast to the similar decay $B^+ \to \eta K^{*+}$ whose QCD penguin contributions appear with opposing signs in the expression for the decay amplitude and hence the predicted branching ratio is substantially higher than the mode studied here. Lipkin [15] predicts that the branching ratios of $B^+ \to \eta K^{*+}$ and $B^+ \to \eta' K^{*+}$ will differ by a factor of around six (no error quoted) if the more sophisticated treatment of $\eta - \eta'$ mixing is used which allows contributions from OZI-suppressed penguin diagrams to be included. If the standard mixing is used, and OZI-terms neglected, the ratio between the branching ratios rises to 33. The branching ratios are therefore significantly affected by the inclusion of OZI-suppressed penguin terms since these sum constructively in the Hamiltonian. Electroweak penguin terms which interfere constructively with other penguin contributions are also expected to contribute to the small branching ratio.

1.4.2 Effective Hamiltonian

The SU(3) approach is useful in generating insights into the importance of each diagrammatic contribution to a decay, or a series of decays. To arrive at a more precise estimate, we must account for QCD corrections. However, a difficulty arises in that there are contributions to the decay amplitude from processes at vastly different energy scales, those from electroweak processes (mediated by the W boson) and those from gluonic processes (mediated by the strong force). This is treated in the Effective Hamiltonian by using an operator product expansion to generate the low energy Hamiltonian from a sum of local operators and Wilson coefficients. This allows the separation of terms arising from gluonic and electroweak processes. The high energy QCD corrections are treated perturbatively due to the short distance they operate over, while the low energy contributions are treated non-perturbatively.

The sum of these components is the effective Hamiltonian. The form of the effective Hamiltonian for the $B^+ \to \eta' K^{*+}$ decay is shown below [19]:

$$\mathcal{H} = \frac{G_F}{\sqrt{2}} \sum_{n=u,c} V_{nb} V_{ns}^* (C_1(\mu) Q_1^n + C_2(\mu) Q_2^n + \sum_{i=3}^{10} C_i(\mu) Q_i^n) \tag{1.32}$$

where G_F is the Fermi constant, $C_i(\mu)$ are the QCD Wilson coefficients which are subject to a renormalisation scale μ, and the local operators are given by Q^n. The operators are dependent on the topology of the processes involved, and can be found in [19]. The limitation of the effective Hamiltonian approach is that there are large uncertainties in the coefficients, form factors and the operators used.

1.4.3 QCD factorisation

Factorisation approaches ignore strong rescattering, which limits their usefulness. QCD factorisation attempts to address this issue. The attraction of a more comprehensive approach is that it ought to be more accurate, giving a more comprehensive account of decay topologies in making predictions of amplitudes and branching fractions. The cost of this more sophisticated approach lies in the fact that many more inputs are required, and some of the quantities may not be known with a great deal of accuracy. This is of particular relevance to the decay channel studied here, as there are additional significant (dominant, even) uncertainties for modes involving an η or η' due to the mixing angle, gluon components in the η and η' wave functions, and an annihilation term in the $B \to \eta^{(\prime)}$ semileptonic form factors. The mass of the strange quark is also a source of considerable uncertainty. The errors in the predicted branching fractions for the mode of interest easily dwarf the actual value, as seen in Table 1.2.

The aim of QCD factorisation is to explicitly calculate the matrix elements of the weak Hamiltonian discussed above, which relate the initial and final states in the decay via transition operators [26, 27]. The matrix elements can be calculated in the heavy quark limit for some two-body final states $M_1 M_2$, (where M_1 contains the spectator quark) and the method reduces the uncertainties associated with the general effective Hamiltonian method.

The QCD factorisation approach requires many input parameters derived from both experiment and theory, such as CKM matrix elements, quark masses and the strong coupling constant in addition to transition form factors and meson decay constants. The $\eta - \eta'$ mixing angle is also relevant in this decay channel and the dominant source of errors.

Other theoretical treatments

Other theoretical approaches involving QCD are also employed. Perturbative QCD (pQCD) was developed in response to divergences that occur within QCD when trying to analyse systems with light, high energy particles. These divergences arise either as a result of energetic quarks interacting with other particles travelling in the same direction (collinear divergences) or due to some particles involved having much less energy than others (infrared divergences) [28–30].

pQCD uses three distinct energy scales, and computes the decay amplitude by factorising it into an expression which is a function of hard scattering kernels, form factors, wave functions and Wilson coefficients. The predictions require experimental input to describe the wave functions, after which predictions can be made regarding decay amplitudes. This approach relies on experimental measurements as inputs to the model and in the cases where such measurements may not exist, uncertainties on predicted quantities may be large.

Chapter 2

The *BABAR* detector in the PEP-II collider

2.1 The PEP-II B factory at SLAC

The PEP-II collider at the Stanford Linear Accelerator Centre (SLAC) was designed and built to produce B mesons. It is an asymmetric e^+e^- collider with a centre of mass energy of 10.58 GeV which corresponds to the $\Upsilon(4S)$ resonance. This resonance is used since the $\Upsilon(4S)$ decays almost exclusively to $B^0\bar{B}^0$ and B^+B^-, making it a clean source of B mesons. The asymmetry of the colliding beams (the electrons have an energy of 9 GeV, the positrons 3.1 GeV) results in the production of B mesons with enough momentum to allow the two B meson decay points to be differentiated through the measurement of decay lengths. This would be impossible in a symmetric colliding system due to the low kinetic energy with which B mesons are produced since the two B masses combined are just less than the $\Upsilon(4S)$ resonance. The asymmetry means that the mesons have appreciable kinetic energy in the lab frame, allowing the mesons' decay vertices to be accurately determined. The asymmetry of the colliding beams results in a Lorentz boost in the direction of the beam axis of $(9 - 3.1)/10.58 \approx 0.56$. PEP-II is described in detail in [31].

To use PEP-II to study CP violation in B meson decays, and to discover rare B meson

decays a high luminosity is required, since the branching fractions involved are often small ($\mathcal{O}(10^{-6})$). PEP-II has a design luminosity of $6\,\text{x}10^{33}\,\text{cm}^{-2}\text{s}^{-1}$, although produced a peak luminosity of over $9\,\text{x}10^{33}\,\text{cm}^{-2}\text{s}^{-1}$ during data-taking Run 4 and a peak of $12\,\text{x}10^{33}\,\text{cm}^{-2}\text{s}^{-1}$ during Run 6.

A diagram of PEP-II and *BABAR* at SLAC is shown in Figure 2.1.

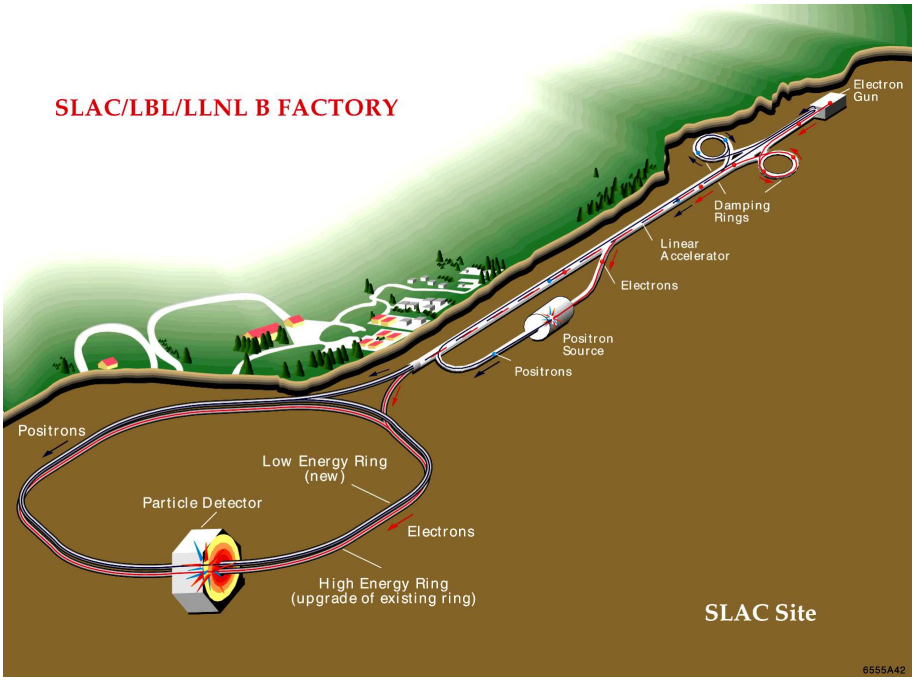

Figure 2.1: PEP-II and *BABAR* at SLAC.

The process for producing collisions is as follows. The electrons are produced by an electron gun, accelerated to high energy by the Linac and injected into the high energy ring (HER). A portion of the produced electrons are fired into a fixed target, producing positrons which are then fed through the accelerator and into the low energy ring (LER), which runs in the opposite direction to the HER. The two beams are brought together at the interaction region by first focussing with quadrupole magnets, and then using a strong dipole field to bring the beams together and to separate them after the interaction point to avoid collisions outside the desired region.

It is desirable to reduce the machine backgrounds to prevent data loss due to increased occupancy, and radiation damage to the subdetectors around the interaction region. The

main sources of background are synchrotron related, injection related, and that caused by lost particles. Synchrotron radiation backgrounds are an artefact of the strong focussing magnets, and to reduce the amount of these photons entering the detector, components have been designed with this problem in mind. The use of copper masks also helps to minimise this. The lost particle backgrounds are caused by the imperfect vacuum in the beampipe where interactions between gas and the beam produce unwanted events in the detector. Injection-related backgrounds are severe and can damage the detector. The method of injection was changed from a system where the rings were filled at once, to a trickle-injection system prior to Run 4. In the older configuration, sensitive subdetectors were powered down during injection process which took around 5 minutes; the beams were stabilised and then the detector was fully powered up and data taking began for a period of around half an hour, during which the current in the ring diminished, and the whole process was repeated. The initial injection phase created large backgrounds as the bunches through the whole ring were unstable at the same time. With the trickle-injection system, single bunches are fed in to the ring at regular intervals. This configuration has several advantages. The beams can be stabilised while trickle injection is occurring, allowing data taking to occur (albeit with a veto applied for the period immediately after injection). Since the beams are stable over much longer periods, the beam conditions could be further optimised for data taking. Also, the time spent taking data was increased since the single injected bunches could be stabilised more easily.

2.2 The *BABAR* detector

The *BABAR* detector has been designed specifically to study CP violating and other rare B meson decays. An end view of the detector is shown in Figure 2.2. It comprises several different elements, namely a silicon vertex tracker (SVT), a drift chamber (DCH), an electromagnetic calorimeter (EMC) and a system used for particle identification (PID) comprising a Čerenkov detector (DIRC) and an instrumented flux return (IFR) for μ/K_L^0 identification. A 1.5T uniform axial magnetic field is generated by a solenoid which enables charged par-

ticles' momentum to be determined. The detector is found around one of the interaction points of the e^+ and e^- beams at PEP-II.

The acceptance of the detector must be maximised in the centre of mass frame. Due to the asymmetric nature of the PEP-II collider, this requires an asymmetric detector. The centre of the detector is shifted along the boost direction from the interaction point of the e^+e^- beams, and also the EMC has an endcap in the direction of the boost only.

The detector must have excellent vertex resolution since the resolution of the difference in B meson decay time is dependent on the accuracy of the decay vertices.

Particle tracking must be accurate over a transverse momentum range of $60\,\mathrm{MeV/c} < p_t < 4\,\mathrm{GeV/c}$, and accurate PID of e, μ, π, K and p is required over a broad kinematic range. The detector must also be able to detect γ and π^0 particles over an energy range of $20\,\mathrm{MeV} < \mathrm{E} < 4\,\mathrm{GeV}$. The *BABAR* detector is described in detail in [32], and this is summarised below.

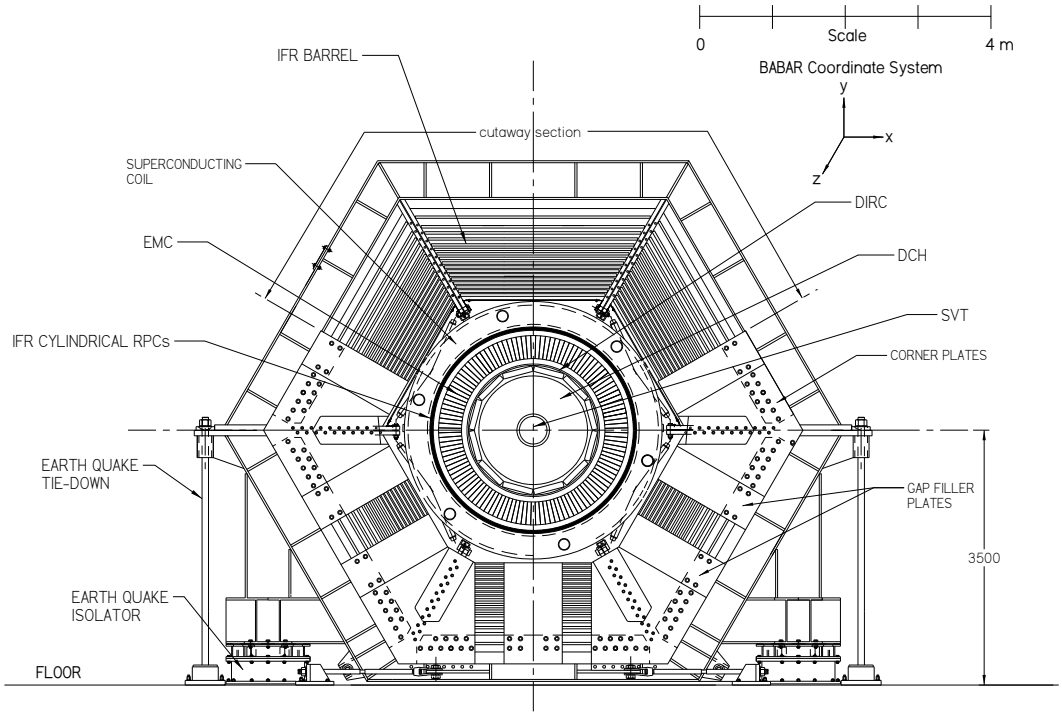

Figure 2.2: *BABAR* detector end view.

2.2.1 The silicon vertex tracker (SVT)

Physics requirements

The task of the SVT is to measure accurately the positions of the two B meson decay vertices. Also, it provides data required to reconstruct charged particle trajectories for particles with momentum of less than 120 MeV/c since these particles do not have sufficient energy to reach the DCH. The best measurements of vertex impact parameters come from the innermost layers, since these are the closest to the interaction point. The limiting factor in these measurements comes from scattering in the beampipe and within the silicon of the detector. The minimum spatial resolution required to study CP violation using the decay length difference is 80 μm. The solid angle coverage of the detector needs to be as high as possible.

Specifications and design of the SVT

The SVT comprises 0.96 m^2 of active detecting area split into five double sided concentric cylindrical shells arranged as shown in Figure 2.3.

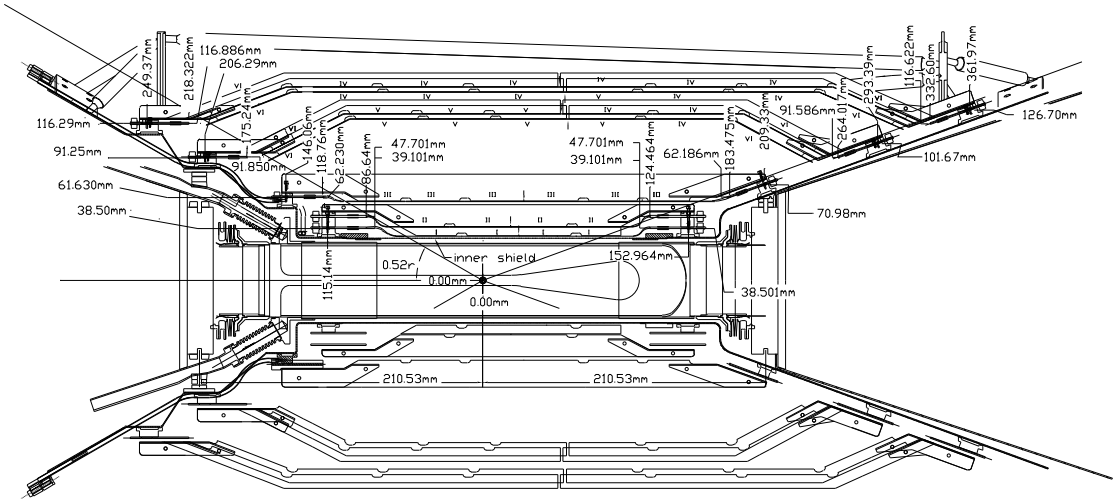

Figure 2.3: Schematic diagram of SVT.

The silicon strip sensors mounted on the inner and outer parts of each shell have the strips oriented orthogonally, to provide 2-D position data. The outer strips of the shells are oriented parallel to the beam direction. The inner three shells are close to the beam pipe and are responsible primarily for measuring initial track parameters, while the outer two layers are set close to the DCH to achieve the best pattern recognition and low transverse momentum (p_t) tracking. The distance between the inner three and outer two shells allow a better measurement of particle tracks. The outer two layers have an arch shape for three reasons: to reduce silicon use; to reduce the amount of material in front of the drift chamber and finally to increase the angle of incidence for particles on the outer edges of the acceptance. The maximum resolution achievable with the SVT is limited by multiple scattering in the beam pipe and the silicon itself. It is for this reason that the best resolution is realised with the innermost shells, shown in Figure 2.4.

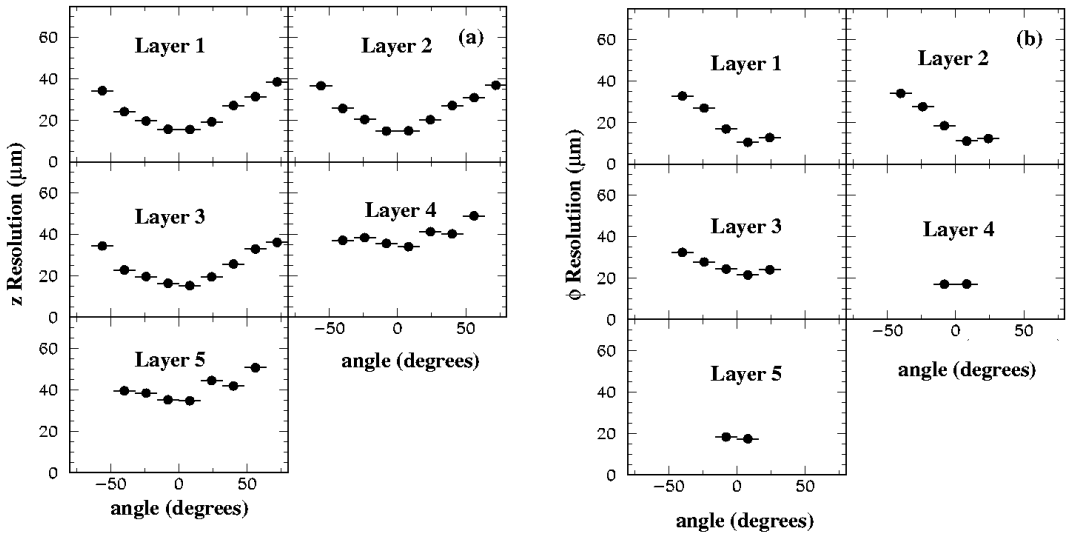

Figure 2.4: Resolution of SVT in z (left plot) and ϕ as function of track angle.

It is clearly desirable to maximise the acceptance of the detector but due to the positioning of essential mechanical, electronic, electromagnetic and optical components, complete coverage is not possible. The polar angle acceptance (θ, measured from the beam axis, z) is restricted to the range $20.1^o < \theta < 150.2^o$. Complete coverage is achieved in azimuthal angles by overlapping the detectors in one of two ways: tilting the modules comprising the innermost

three layers by 5^o in the ϕ direction; dividing each of the two outermost layers into two sublayers, with each sublayer placed at a slightly different radius, as shown in Figure 2.5.

The SVT is vulnerable to radiation damage from synchrotron backgrounds, with the detectors in the horizontal plane being most at risk. There is a safety system which trips if conditions become dangerous for the detector. The SVT is mounted inside a support tube connected to the final beam magnets, in contrast to the other subdetectors which are attached to outer components of the detector.

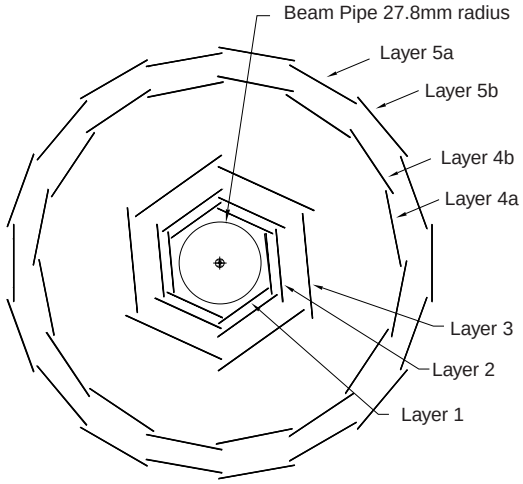

Figure 2.5: Transverse section of SVT.

2.2.2 The drift chamber (DCH)

Physics requirements

The DCH is the principal device by which the momenta and positions of charged particles are recorded in the detector, and complements the charged track direction and impact parameter measurements recorded by the SVT. The DCH is crucial to extrapolating the charged tracks to other components in the *BABAR* detector such as the DIRC, EMC and IFR. It also provides PID by measuring ionisation loss (dE/dx) for low momentum particles which are not registered by the DIRC. The range of acceptance of the DCH is $17.9^o < \theta < 162.1^o$ (this is limited by PEP-II beam components) and in the extreme forward and backward directions

the DCH is the only means of identifying particles of different masses. A dE/dx resolution of 7% is required to discriminate between π and K of momentum up to 700 MeV/c, while for particles with p_t of greater than 1 GeV/c the resolution required is 0.3%.

It is crucial that the DCH provides high resolution data in order to aid particle identification and match tracks to the DIRC. One of the most significant factors which limits the performance of the DCH and other detectors further from the interaction region is multiple scattering. In order to minimise this effect, the amount of material in the DCH must be minimised.

Specifications and design of the DCH

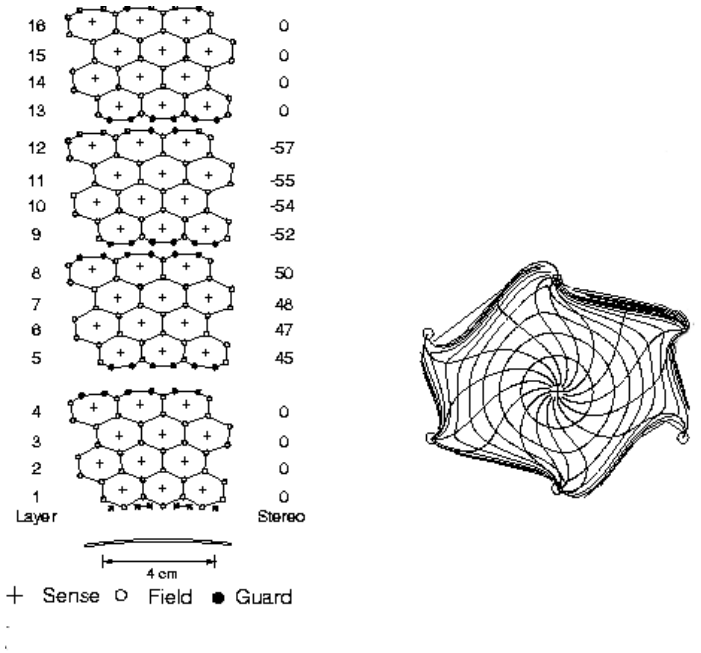

Figure 2.6: Layout of cells in innermost layers of DCH (left) and 100 ns isochrones in a typical drift chamber cell.

The DCH is a 2.8m long annulus with an inner radius of 0.24 metres (the SVT is located in the centre) and outer radius of 1.6m. As with the SVT, the DCH is not centred on the IP along the beam axis in order to provide better coverage in the forward direction. Also, the

electronics are mounted on the rear endplate of the DCH, and the forward endcap is half the thickness of the rear (12mm & 24mm respectively) in order to minimise the material in front of the EMC endcap in the forward direction. The inner wall of the annulus is constructed from 1mm beryllium ($0.28\%\,\mathrm{X}_o$) and the outer wall from two layers of carbon-fibre mounted on a honeycomb core ($1.5\%\,\mathrm{X}_o$). The wires in the DCH are strung between the aluminium endcaps and there are two types of wire. The wires used for sensing are 20μm diameter tungsten-rhenium. A potential difference of 1,960 V exists between these wires and the other, field wires, which are electrically grounded. These wires are arranged such that a hexagonal area (cell) of ($\approx 2\mathrm{cm}^2$) is defined around the wire when looking along the wire axis, shown in Figure 2.6.

The DCH consists of 7,104 cells arranged into 10 superlayers with each superlayer comprising 4 layers. Successive superlayers' wires have different orientations, cycling between axial (A) and stereo (U,V) superlayers. The stereo superlayer wires vary in angle between 40 and 70 mrad relative to the axial wires in order to allow longitudinal position data to be gathered. The DCH is filled with a 4:1 mixture of helium/isobutane as this provides good spatial and dE/dx resolution while providing a low drift time and minimising multiple scattering due to its low mass. Additional 'guard' and 'clearing' wires are added at the boundaries of superlayers and the extremities of the detector in order to ensure consistent performance across all cells and to collect charges created via interactions in the walls of the detector.

The cell arrangement allows up to 40 spatial and ionisation loss measurements for particles with transverse momenta p_t of $> 180\,\mathrm{MeV/c}$. The reconstruction of tracks uses data gathered from both the DCH and the SVT. The tracks are parameterised by 5 quantities, namely: d_o, the distance of closest approach of the track to the origin in the $x-y$ plane; z_o, the distance of closest approach of the track to the origin along the z-axis; ϕ_o, the azimuthal angle of the track; λ is the dip angle relative to the transverse $x-y$ plane and $\omega = 1/p_t$ defined as the track curvature. These parameters are used with a fitting algorithm to reconstruct the tracks. Figure 2.7 shows dE/dx measurements as a function of momentum and the solid lines are the Bethe-Bloch curves for each species of particle, and a performance plot of the transverse momentum resolution obtained by measurements made by the SVT and DCH.

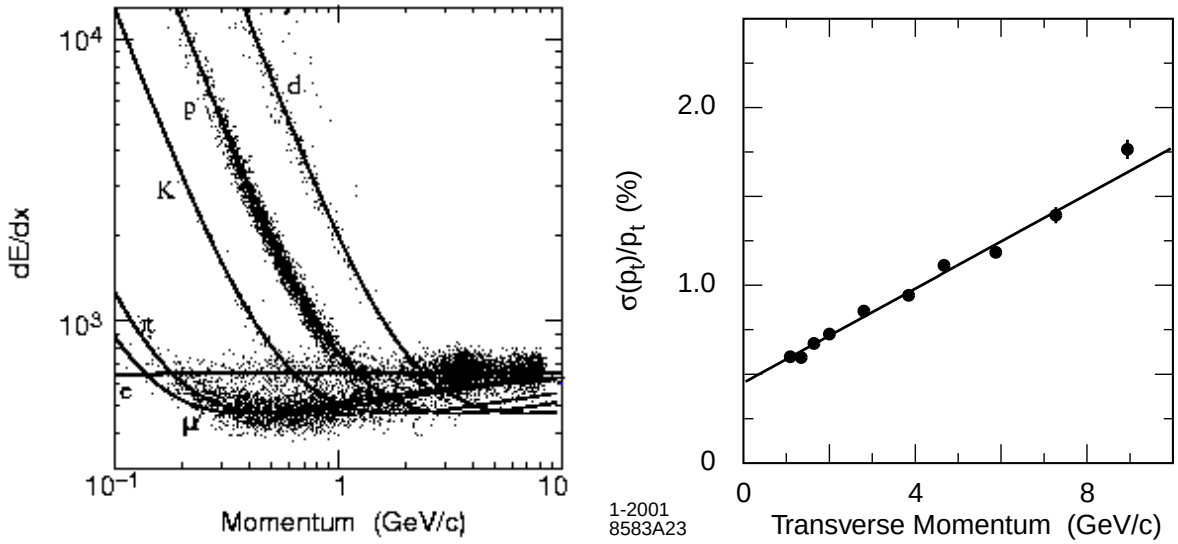

Figure 2.7: On the left, dE/dx measurements as a function of momentum in the DCH, with Bethe-Bloch parametrisation curves overlaid. Bhabha(e^+e^-) scattering is evident in the high momentum regions ($3-8$ GeV/c). On the right, a plot of transverse momentum resolution determined from cosmic ray muon studies, which is primarily dependent on the DCH.

Kaons and pions can be discriminated from each other up to momenta of around 700 MeV/c.

2.2.3 The detector of internally reflected Čerenkov light (DIRC)

Physics requirements

The primary function of the DIRC (Detector of Internally Reflected Čerenkov light) is particle identification (PID). In the momentum range 1.7 to 4.2 MeV/c, excellent π and K separation is required for studies of rare B decays and the DIRC has therefore been designed to provide a $4\,\sigma$ separation across this range. No π/K separation occurs at 1.1 GeV/c. For particles with momenta less than 700 MeV/c, $K\pi$ PID comes from ionisation loss data gathered in the SVT and DCH. Above 700 MeV/c the DIRC must provide this. Again, the amount of material present in the detector must be kept as low as possible to compromise

the performance of the other subdetectors as little as possible. The spatial coverage of the detector is maximised.

Design of the DIRC

Čerenkov radiation will be emitted by a particle which has a speed greater than the local speed of light when $\beta > 1/n$, where β is the speed as a fraction of the speed of light and n is the refractive index of the material. The radiation is emitted in a cone with opening angle $\cos\theta_c = 1/(n\beta)$. The DIRC is different from most Čerenkov detectors in that it relies on total internal reflection. The radiation is emitted when a particle passes through one of 144 quartz bars (radiators). Each bar is 1.7cm thick, 3.5cm wide and 4.9m long. Quartz is used since it has many desirable optical properties such as low chromatic dispersion, good radiation hardness and it can be polished to a high degree. The radiation produced travels along a quartz bar to the detectors at the backward end of the detector. A schematic diagram of the DIRC is shown in Figure 2.8. The quartz preserves the opening angle θ_c of the radiation when it is reflected at the boundaries of the bar, and this image is detected when the radiation exits the bar into a standoff box filled with water which is surrounded by an array of 10,752 photomultiplier tubes. These detect radiation in the near UV and visible ranges. A silica wedge attached to the end of the bars between the bars and the standoff box allows photons emerging from the bars at large angles to be reflected back towards the central axis, allowing photons to be detected which would otherwise be lost without a larger array of photomultiplier tubes. Additionally, radiation travelling in the forward direction is reflected by mirrors at the bar ends to the backward end to be detected. The detected image is several conic sections, where the cone's opening angle θ_c is one of the Čerenkov coordinates, but modified by the refractive properties of the silica and water media through which the photon has travelled. The spatial and timing data from the photomultiplier tubes are used with track data from the bars to reconstruct an event via a maximum likelihood fit. The timing data are also used with the track information to reject beam background events.

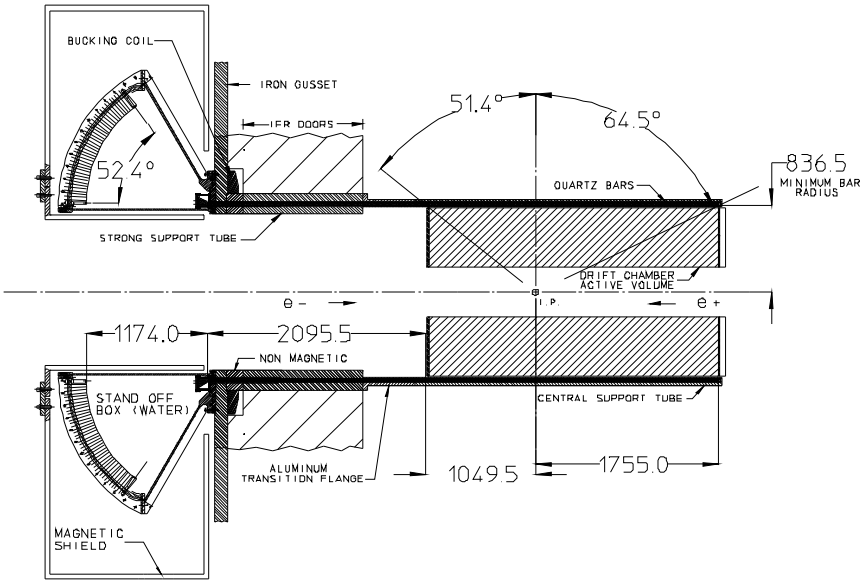

Figure 2.8: Side view of the DIRC.

2.2.4 The electromagnetic calorimeter (EMC)

Physics requirements

The function of the EMC is to detect and measure photon energies from decay processes, the most significant of which for this analysis are η and π^0 decays. It is also used to identify electrons for B flavour tagging in conjunction with data from the DCH. In order to fulfil these roles, the EMC needs to be effective over a large range of energies, from 20 MeV to 9 GeV. The upper bound comes from the need to measure QED processes required for calibration (Bhabha scattering) and luminosity, and the lower from the need for reconstruction of events containing multiple π^0 and η particles.

Design and construction of the EMC

The detection medium of the EMC is crystalline CsI(Tl) with silicon photodiode readout chips. There are 6,580 crystals arranged in a cylindrical shell with a conical endcap in the forward direction, as shown in Figure 2.9. The EMC provides full acceptance in azimuthal angle and polar angle coverage extends between 15.8^o and 141.8^o. The combined coverage

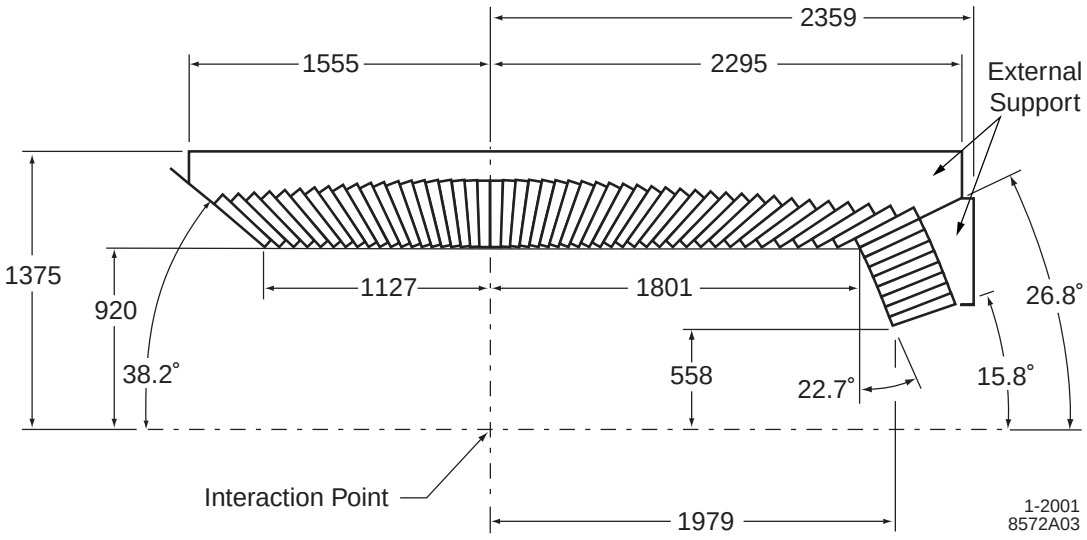

Figure 2.9: Schematic diagram of the top half of the EMC. The high energy beam enters from the left hand side.

amounts to 90 % of the solid angle in the centre of mass system. The EMC is designed to be a total absorption calorimeter and it is for this reason that the depth of the crystal shell increases towards the forward direction. In addition, the crystals are arranged such that they project out from a point slightly away from the interaction region which minimises losses due to gaps between them, estimated to be $\approx$ 2.5 %. The crystals are wrapped in a white reflecting material (Tyvek) to minimise transmission losses and then isolated electromagnetically with a foil Faraday shield and electrically using a mylar layer. The barrel and outer rings of the endcap have less than $0.3 - 0.6\,\mathrm{X}_0$ in front of the crystals, and the crystals themselves have a radiation length of 1.85 cm.

A typical event involves an electromagnetic shower which is detected in several adjacent crystals. The reconstruction algorithm applies a minimum energy cut (20 MeV) to the crystals in the cluster and then the crystal with the highest energy deposit is identified. The exact positions of the energy maxima in a cluster are determined via iteration and the cluster position can then be compared to charged track data to identify a candidate particle. If no charged track can be projected to the position of the energy maxima in the EMC crystal, then a neutral candidate particle is assumed.

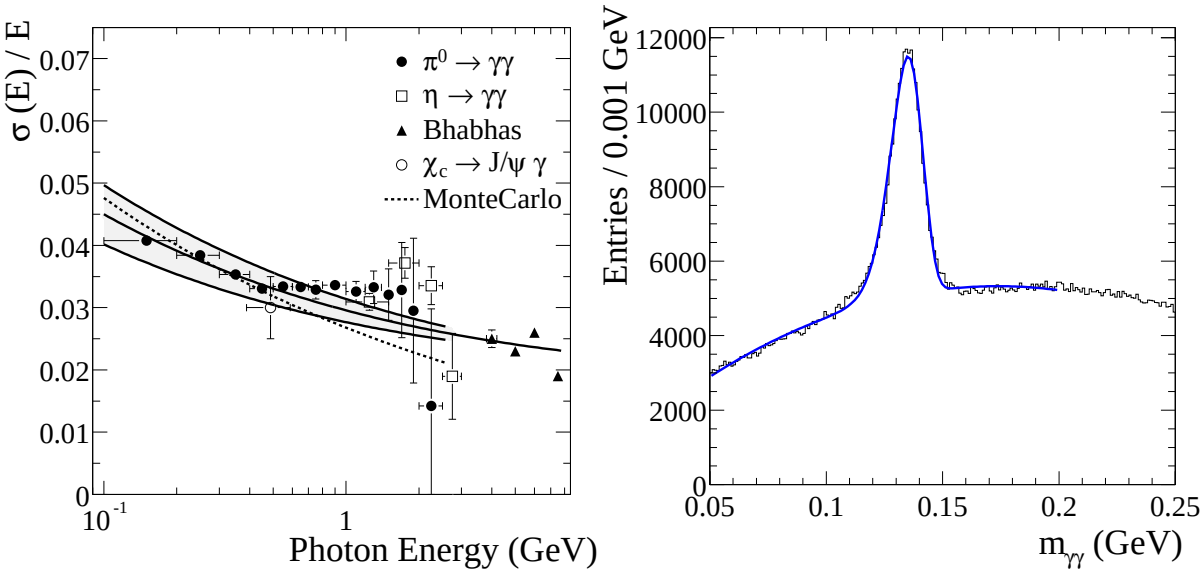

Figure 2.10: Plots showing aspects of EMC performance. On the left, energy resolution of the EMC for photons for various processes, including π^0 decays. The fitted solid lines are an empirical parameterisation describing the energy resolution of a homogeneous crystal calorimeter, with rms errors plotted above and below, bounding the shaded area. On the right, a plot of the invariant mass of two photons in $B\bar{B}$ events. The solid line is a fit to the data, which agrees well with MC simulations.

The data from the clusters are used in combination with those from other subdetectors to discriminate between candidate particles. For example, electrons and charged hadrons are discriminated via the momentum and energy of the shower, track momentum and data regarding energy loss in the DIRC and Čerenkov angle. Misidentification of π^+ and e is dependent on selection criteria on the track momentum, but typically is under 0.5%. Photon energy resolution from π^0 decays is of particular relevance to this analysis, as is π^0 reconstruction from two photons. The performance plots for each are shown in Figure 2.10. These agree well with or exceed expected performance from MC simulations.

2.2.5 The instrumented flux return (IFR)

Physics requirements

A flux return is employed to optimise the magnetic field conditions, and to localise the field. This passive role can be achieved while also allowing the flux return a role in PID. The function of the IFR is to identify and reconstruct muons and neutral hadrons e.g. K_L^0. Muon identification is important for flavour tagging neutral B mesons involved in semi-leptonic decays, and K_L^0 mesons are produced in some exclusive B decays to CP eigenstates of particular interest. The discrimination between muons and neutral hadrons requires information from other parts of the detector, namely the SVT/EMC/DIRC for muons and EMC for neutral hadrons.

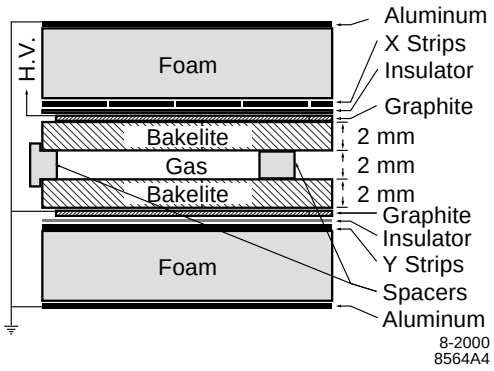

Figure 2.11: Cross-section of an RPC in the IFR employed in Runs 1-4.

Design of IFR used in Runs 1-4

This comprises 806 resistive plate chambers (RPCs) sited within the magnet assembly on the detector. A cross section of an RPC is shown in Figure 2.11. The gas chamber is filled with non-flammable mixture of $\approx 55\%$ argon, $\approx 40\%$ freon and $\approx 5\%$ isobutane. A potential difference of 8 kV is maintained across the chamber by connecting the graphite-coated Bakelite to the power source. The RPCs detect streamers from ionising particles, and the event is read out using capacitive readout strips which could achieve a position resolution

of a few mm. The RPCs are arranged in a central barrel with two endcaps (see Figure 2.12) which ensure coverage from 300 mrad in the forward direction to 400 mrad in the backward direction. There are 19 layers of RPCs in the barrel and 18 in each end cap. The forward endcap was upgraded to better RPCs prior to Run 4.

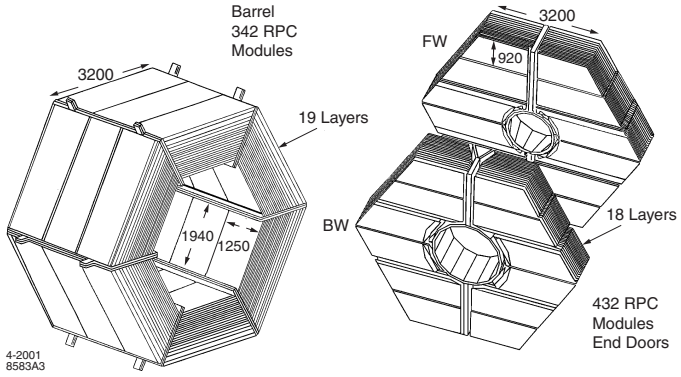

Figure 2.12: Overview of the IFR employed in Runs 1-4.

Modifications to the IFR carried out for Runs 5 & 6

Due to degradation in the efficiency of the IFR over the operation of the experiment, the IFR was modified by the inclusion of Limited Streamer Tubes (LSTs) in the barrel section. These were installed during the two periods of downtime between Runs 4, 5 and 6. The LST system was implemented due to its proven capability in other experiments (eg ZEUS at DESY in Hamburg), and the urgent need to restore identification efficiency and accuracy. In common with the system it replaced, it uses gas filled chambers. The RPCs in the barrel were removed in two stages, and replaced with LSTs which consist of eight cells of 1 cm section. A silver wire runs down the centre of each cell with spacers every 50 cm or so to ensure the wire is centred properly, and the tubes are filled with a non-flammable CO_2 gas mixture. The readout electronics and gas inlet valves are mounted at one end. Each cell can provide a one dimensional co-ordinate, and strip planes are used in order to exactly specify the co-ordinates of a passing particle. The operating voltage is typically 4.7 kV, and the signals on the wire are around $150 - 200$ mV. The LSTs were installed and tested before the start of Run 5 in 2005, and can be seen in Figure 2.13. The barrel replacement program

saw LSTs installed in the top and bottom sections prior to Run 5, and the four side sections were modified before Run 6. The endcaps were unmodified since the rearward endcap covers a small solid angle and the upgrade was not deemed worthwhile, and the forward endcap already had upgraded RPCs fitted prior to Run 4.

Figure 2.13: A typical LST board. The eight cells can be seen, as can the gas inlet valves.

2.2.6 The trigger

Physics requirements

The trigger is the means by which events of interest are selected. The system must be able to accurately reject background events with upwards of 95% efficiency and be robust enough to deal with the physical effects of high background levels on electronics. It must also have a high degree of redundancy with respect to failing electronics and also perform sufficiently well that its own operation does not limit data taking significantly. The trigger is the means by which the collision event rate can be reduced to an acceptable rate for mass storage, without losing efficiency for B decays.

The trigger's functionality is implemented in two levels; a "Level 1" (L1) system which is purely based on hardware output, and a "Level 3" (L3) system which takes L1 input, and applies software processing to reduce the acceptance rate further.

Design of L1 trigger

The L1 trigger is firstly comprised of outputs from processors in three subdetectors: The DCH, EMC and IFR. The design of the system is shown in Figure 2.14. Each of these processors sends data regarding the presence of particle(s) to the Global L1 trigger. The Global L1 trigger receives data from the three processors and if selection criteria within the Global L1 trigger are satisfied, implemented via lookup tables and logic, an instruction to read out the event is issued, and the data are written to disk.

The L1 trigger has an output rate of around 2kHz under normal operation.

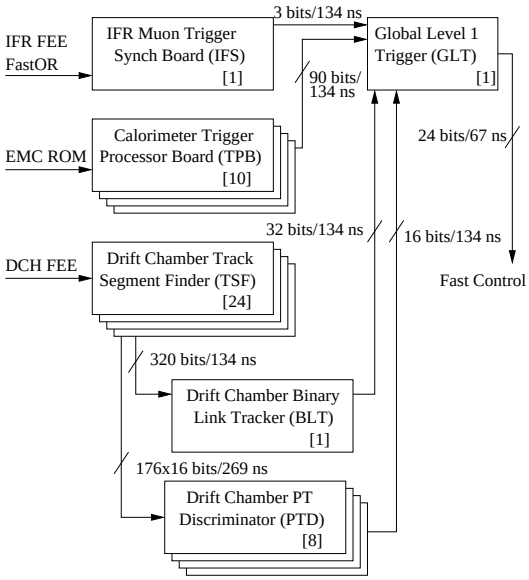

Figure 2.14: An outline of the L1 trigger, showing the three subdetector sources of data for the Global L1 trigger. Also shown are signal sizes and module numbers(brackets).

The DCH trigger receives time data from each of the 7,104 DCH cells. This is used to produce a map of candidate tracks via the Track Segment Finder (TSF) and Binary Link Tracker (BLT) modules. The TSF looks for straight tracks originating from the interaction point. The track segments identified by the TSF are passed to the BLT, which links the segments into complete tracks. A transverse momentum cut is applied in a third set of modules before the data are sent to the Global L1 trigger

The EMC trigger is implemented by dividing the subdetector into 280 towers of between

19 and 24 crystals depending on whether the tower is in the barrel or the endcap. Each tower sends data for any crystal which shows an energy deposition of at least 20 MeV to the EMC trigger, which uses lookup tables for threshold energies for various types of event before sending the results to the Global L1 trigger.

Prior to Run 4 the L1 trigger was upgraded via the introduction of a cut on the z-vertex using information from the DCH. Background events are found fairly evenly distributed in z, whereas the events of interest are tightly clustered around $z = 0$. A cut on z is an effective means of reducing the L1 acceptance rate for background events.

The IFR trigger is used to highlight $\mu^+\mu^-$ and cosmic rays, which are useful for calibration.

Design of L3 trigger

The L3 trigger has access to the output from the L1 trigger in addition to comprehensive event and tracking data, and uses this to further refine event selection. Due to limited tracking available to the L1 trigger, the L3 trigger is able to reject a much greater proportion of background events which do not originate from the interaction point. In addition it can be used to prescale Bhabha events to 1 Hz for calibration purposes via selection and classification criteria, and filters in the EMC and DCH. This is of critical importance given the outputs of the DCH and EMC are predominantly events which should be rejected. The output of the L3 trigger is limited to 120 Hz.

The L3 trigger runs within the Online Event Processing (OEP) framework, with the OEP providing events to the L3 trigger and logging the output. The L3 trigger follows a three-phase process which classifies events, runs scripts on the classifications in order to assess whether they pass selection criteria and flag the event and then forms the output based on the flags from phase two. The third phase can veto events flagged in the second phase so backgrounds can be reduced further.

Chapter 3

Data-sets and event selection

3.1 Introduction

The analysis presented here is principally a branching fraction study of the charmless charged decay $B^+ \to \eta' K^{*+}_{K^+\pi^0}$. Two modes are studied because the η' can either decay to $\rho\gamma$ or $\eta\pi\pi$. The charge conjugates are also included. The final combined fit to $B^+ \to \eta' K^*$ presented in Chapter 6 is composed of six modes, the additional four being two charged and two neutral modes:

- $B^+ \to \eta'_{\eta\pi\pi} K^*_{K^0\pi^+}$
- $B^+ \to \eta'_{\rho\gamma} K^*_{K^0\pi^+}$
- $B^0 \to \eta'_{\eta\pi\pi} K^{*0}$
- $B^0 \to \eta'_{\rho\gamma} K^{*0}$.

In the two neutral modes, the K^{*0} decays to $K_s\pi^0$.

To obtain meaningful results from the raw data collected by the *BABAR* detector, the successful separation of signal events from unwanted background events must occur. It is therefore necessary to build the raw data from interactions in the various subdetectors into a coherent picture of a particular event. This reconstructed event may then be analysed against

some kinematic and topological criteria to determine whether to classify it as a signal or background event. There are competing pressures involved in this process, since if selection criteria are applied too tightly the risk of rejecting signal events increases; if cuts are applied too loosely, unwanted backgrounds may swamp the signal. The signal events have branching ratios of the order of 10^{-6}, and so the task of separating signal events from the background is a delicate one. In general, cuts are applied loosely in order to preserve the signal as efficiently as possible, and the results are extracted from the reconstructed event data via the maximum likelihood fit described in Chapter 4. To characterise backgrounds precisely, data from large m_{ES} or ΔE sidebands are used, avoiding the signal region.

The quantities that are reconstructed include the mass, resonant mass, energy and helicity angle of the B and/or daughter particle(s). Topological quantities such as angles of decay, thrust, the location of decay vertices and angular distribution of the decay event are used. The skims detailed below use this reconstructed event data to search for events which loosely fit the final states of interest.

Much of the dominant background is easily identifiable. The continuum ($q\bar{q}$) background, discussed further in Section 4.2, is topologically very different to signal events, having a jet-like distribution. It can be effectively rejected by constructing selection criteria to exploit the jet-like topology. Other background sources are $B\bar{B}$ events which closely resemble the final states of interest. The treatment of such events is detailed in Section 4.2. Some of these are accounted for by quantifying the individual final states and including a weighted, exclusive fit component, e.g. charmless $B\bar{B}$ decays, and the remaining $b \to c$ charmed backgrounds are accounted for, e.g. in the treatment of the $q\bar{q}$ backgrounds. Finally, background from processes such as lepton pair production and other QED effects can be removed by imposing a minimum track requirement on reconstructed events.

3.2 Data-sets

The data used in the final analysis comprise the data-set gathered over runs 1-4 between 1999 and 2004. This contains some $210.5\,\mathrm{fb}^{-1}$ collected on the $\Upsilon(4S)$ which contains 231.8 million

B meson decays.

The signal Monte Carlo (MC) data-sets, derived from GEANT4 simulations [33] of the *BABAR* detector, comprise 124,000 and 125,000 events for $\eta\pi\pi$ and $\rho\gamma$ respectively. The signal MC is used to test the robustness of the analysis technique and fitting procedure, before the analysis technique and fitting procedure is applied to the full on-peak data-set. Thus the analysis is performed "blind", which removes the possibility that biases are introduced by tuning event selection criteria to the data. The final fit to the on-peak data-set is performed only once the fitting procedure has been shown to be sound.

For the $B\bar{B}$ background studies described in Sections 4.2 and 4.5, which quantify the effects of any other known B decay modes being mistaken for the decay channels of interest, a sample of 670 million $B\bar{B}$ MC events is used. For $q\bar{q}$ studies, data from the m_{ES} sideband are used, defined by $m_{ES} < 5.27\,\text{GeV}$ for all variables except m_{ES} which uses data from the ΔE sidebands defined as $|\Delta E| > 0.1\,\text{GeV}$. This allows the signal region to be avoided in order to be able to describe the background more precisely.

3.3 Skims

"Skims" are used as a means of imposing light preselection criteria on the data-set which can remove many of the easily-identifiable background events with negligible effect on signal events. This reduces the amount of more intensive processing required when imposing the detailed selection criteria described later. Two "skims" are used in this analysis, depending on the decay channel of the η'. The "skims" use the reconstructed topological data for each event in order to search for events of interest, which are then further filtered by the use of cuts on invariant masses of candidate particles.

An inclusive skim is used to select events with a final state $\eta' \to \eta\pi\pi$ which assumes two-body kinematics, but without constraint on the other body, since these channels are relatively clean and free from backgrounds. An exclusive skim, which also assumes two-body kinematics, is used to select candidates in the decay channel $\eta' \to \rho^0\gamma$, but since these channels have

substantially higher background contributions, the skim places exclusive constraints on the final states permitted. The skims are detailed below.

3.3.1 InclEta

The skim InclEta looks for and tags the highest momentum η or η' in an event, reconstructing η, η' in the decay channels $\eta \to \gamma\gamma$ and $\eta' \to \eta_{\gamma\gamma}\pi^+\pi^-$.

The η' candidates are reconstructed by combining an η candidate with two charged tracks. A cut on the invariant mass of the η' candidate is performed, constraining it to between 0.9 and $1.01\,\mathrm{GeV}/c^2$.

The skim then selects events if the η or η' has a centre-of-mass momentum between 1.9 and 3.1 GeV/c. This skim selects events from data with fraction 2.67 % and $B\bar{B}$ Monte Carlo with a fraction of 2.41 %.

3.3.2 ExclEtap

The skim ExclEtaP searches for 16 B decays of the form: $B \to \eta' X$ with X being a charmless hadronic state, reconstructing the η' in its decay $\eta' \to \rho^0\gamma$. The ρ^0 candidates are selected by combining two charged tracks to produce a mass within ± 300 MeV of the ρ^o mass. These candidates are then combined with photons and a mass cut is applied to the resultant state, constraining the candidates to $0.9 \leq m^{\eta'}_{\rho\gamma} \leq 1.01\,\mathrm{GeV/c^2}$

The selection cuts applied by this skim are:

- η' centre-of-mass momentum between 1.9 and 3.1 GeV/c,
- B mass greater than $5.15\,\mathrm{GeV/c^2}$,
- B energy minus centre-of-mass beam energy (see Section 3.4) = $\Delta E < 0.3$ GeV,
- momentum of the γ from η: $p_\gamma > 0.050$ GeV/c.

This skim selects signal events with an efficiency of approximately 40 %, depending on the decay channel, with background events passing with an efficiency of a few percent. The details are shown in Table 3.1. The skim assigns Boolean tags to different groups of final states, and one event may have more than one tag assigned to it.

Table 3.1: The sixteen final states which are assigned a Boolean tag with skim ExclEtaP.

Final state	Signal efficiency (%)	$q\bar{q}$ efficiency (%)
$\eta' K_s^0$	48	0.6
$\eta' K^+/\pi^+$	55	2.0
$\eta' K^*$ ($4K^*$ decay modes)	≈ 40	3.8
$\eta' \rho^0/\rho^+$	58/36	< 4.5
η' $\pi^0/\phi/\eta'$ ($2\eta'$ decay modes)	–	2.0
η' ω/η (2η decay modes)	–	1.8

3.3.3 Output

After the skims have been applied, the data-set is composed of events whose final states are generally consistent with those of interest, in this case a quasi two-body decay with products compatible with η' and K^*. The data-set is ready to undergo the second stage of filtering.

3.4 Reconstruction and preliminary cuts

The second stage of filtering is performed to constrain the general final states to be closer to the ones which contain the specific states of interest. This involves track parameters, masses, energies and momenta of the particles of interest, and is where the fine tuning of analysis variables occurs, prior to the fitting procedures which will extract the signal yields and other quantities which we are ultimately interested in. The analysis variables and the

constraints imposed are described below.

ΔE

A B meson candidate is characterised kinematically by the invariant ΔE where

$$\Delta E = E_B^{CM} - E_{Beam}^{CM} \tag{3.1}$$

with E_B^{CM} being the B energy in the centre-of-mass (CM) frame and E_{Beam}^{CM} is the beam energy in the same frame. The beam energy is tuned to the $\Upsilon(4S)$ resonance, which is produced at rest in the CM frame. Each resultant B has half that energy in the same frame since they are produced in pairs. ΔE is Gaussian distributed and centred on zero for signal events.

$\mathbf{m_{ES}}$

The beam energy substituted mass, m_{ES} is defined as:

$$m_{ES} = \sqrt{(E_{Beam}^{CM})^2 - (p_B^*)^2} \tag{3.2}$$

where p_B is the B momentum in the CM frame and E_{Beam}^{CM} is the beam energy in the CM frame. Due to the precision with which the beam energy is known, this formulation of the B mass gives better resolution than that which could be achieved via reconstruction of the B energy.

Resonant masses: m^{η}, $m^{\eta'}$, m^{K^*} and m^{π^0}

The various products are either detected directly (as in the case with kaons) or reconstructed from daughter product tracks and clusters (as with π^0, η and η'). The resonances are reconstructed in the following way:

- A B candidate is formed by combining an η' candidate with a K^*.
- An η and two charged tracks, or a ρ and a photon are combined to form an η' candidate. A mass constraint is applied to the η' candidate to both improve B candidate ΔE resolution and to reduce correlations between ΔE and η' mass.
- Two photons are combined to form a π^0 or an $\eta_{\gamma\gamma}$ candidate. The unconstrained invariant mass is computed for the purposes of their selection, and then the mass is constrained to the PDG value for the parent's composition.
- A K_s^0 or a π^0 is combined with a charged track to form a $K^{*+} \to K_s\pi^+$ or $K^{*+} \to K^+\pi^0$ candidate.

K^* helicity angle, $\cos\theta_H$

The helicity angle of the K^* vector meson, $\cos\theta_H$ is defined as the cosine of the angle between the daughter pion of the resonance and negative of the B momentum in the K^* rest frame. The decay topology has significant consequences for background contributions, and helicity angle is a powerful tool with which to remove background contributions from modes which mimic the final states of interest. When the daughter pion is travelling against the original B momentum it has the lowest possible momentum. This corresponds to a helicity angle of 1. Since neutral pions are easily faked in the calorimeter, the signal yield is vulnerable to contributions from modes which mimic the signal. This is the principal reason for enforcing cuts on the helicity angle of recorded events. The details and explanation of the helicity angle cuts applied to each mode are found in Section 3.5.

Thrust angle, $\cos\theta_T$

The thrust angle is defined as the angle between the B candidate and the rest of the event. It is very effective at discriminating between signal and $q\bar{q}$ events since the topology of the latter is highly jet-like along the direction of quark production, whereas the thrust angle for

B candidate events has no preferred direction. We apply different cuts to $\cos\theta_{\rm T}$ to optimise the signal efficiency in each mode, as shown in Table 3.2.

Fisher discriminant, $\mathcal{F}$

The dominant backgrounds, attributable to $q\bar{q}$ events, have several measurable characteristics which are correlated. Since the fitting method described in Chapter 4 requires independence between input variables, it is not appropriate to put all these quantities to the fit directly. It is possible to construct a composite quantity which can exploit the various correlations however. The Fisher discriminant [34] is a powerful tool for separating $q\bar{q}$ events from signal events by identifying the jet-like topology of the former. Since $q\bar{q}$ backgrounds account for most of the recorded events, it is imperative that they are rejected efficiently. The Fisher is constructed of a weighted linear combination of two angles and two Legendre monomials ($\rm L_0$ and $\rm L_2$). The angles are: $|\cos\theta_C|$, the angle between the signal candidate thrust axis and the beam axis, and θ_B which is the angle between the B candidate momentum and the beam axis. The monomials describe the angular distribution of the momentum flow from the part of the event unrelated to the B thrust axis. The weightings in the Fisher were optimised to provide maximum separation between signal and $q\bar{q}$ events, and the Fisher is used as an input to the maximum likelihood fit described in Chapter 4.

3.5 Details of preliminary cuts applied

- $5.25 \leq \mathrm{m}_{ES} \leq 5.29\,\mathrm{GeV}$,
- $N_{\rm trks} \geq \max[3, N_{\rm tracks\,in\,decay\,mode} + 1]$ (to be able to define a thrust vector for the rest of the event),
- $120 < m^{\pi^0}_{\gamma\gamma} < 150\,\mathrm{MeV/c^2}$,
- $490 < m^{\eta}_{\gamma\gamma} < 600\,\mathrm{MeV/c^2}$,
- $910 < m^{\eta'}_{\eta\pi\pi/\rho\gamma} < 1,000\,\mathrm{MeV/c^2}$,

- $755 < m_{K\pi}^{K^*} < 1,035\,\mathrm{MeV/c^2}$,
- $-4 < \mathcal{F} < 5$,
- daughters of $\pi^0 \to \gamma\gamma$ candidates are required to have $E_\gamma > 30\,\mathrm{MeV}$,
- daughters of $\eta \to \gamma\gamma$ candidates are required to have $E_\gamma > 100\,\mathrm{MeV}$,
- daughters of $\eta \to \rho^0\gamma$ candidates are required to have $E_\gamma > 200\,\mathrm{MeV}$.

The subdecay modes use different cuts for ΔE, $\cos\theta_H$ and $\cos\theta_\mathrm{T}$, detailed in Table 3.2 below.

ΔE is cut asymmetrically around the peak in the distribution. The tighter upper limit is used to reduce $B\bar{B}$ backgrounds, and those modes with a neutral kaon can be cut more tightly due to the better resolution, but not so tightly to risk losing signal events. The low side of ΔE generally shows a long tail which is a result of energy lost between crystals in the EMC and bremsstrahlung. The ΔE distribution is broader, and has more pronounced tails for modes where $K^{*+} \to K^+\pi^0$, due to the π^0 produced.

The helicity angle, $\cos\theta_H$ cut depends on two factors: the charge of the pions, and the η' decay channel. Decay modes with π^0, rather than π^+, are more tightly cut since the calorimeter background is larger, and the potential to misidentify the π^0 candidate is high. A slow moving π^0 candidate, found at $\cos\theta_H = 1$, is easily faked and could be included with the decay mode $\eta' K$ to mimic the final state of interest. In addition, the mode with both a π^0 and the $\eta' \to \rho\gamma$ decay channel is constrained further. A looser cut can be used for charged pions since there is less background. A cut of 0.95 is imposed on these modes to remove the slowest charged pions. All of these cuts reduce the $B\bar{B}$ backgrounds.

The $\cos\theta_\mathrm{T}$ cut is a powerful tool for removing QCD backgrounds, as these events show a strong jet-like topology. Tighter cuts are applied for the $\eta' \to \rho\gamma$ decay channels due to the higher backgrounds in these modes.

Table 3.2: Summary of the ΔE, $\cos\theta_H$ and $\cos\theta_{\rm T}$ cuts applied.

Mode	ΔE(GeV)	$\cos\theta_H$	$\lvert\cos\theta_{\rm T}\rvert$
$\eta'_{\eta\pi\pi}K^{*0}$	$-0.200 < \Delta E < 0.125$	$-0.95 < \cos\theta_H < 1.00$	0.90
$\eta'_{\rho\gamma}K^{*0}$	$-0.200 < \Delta E < 0.125$	$-0.95 < \cos\theta_H < 1.00$	0.75
$\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$	$-0.200 < \Delta E < 0.125$	$-0.95 < \cos\theta_H < 1.00$	0.90
$\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$	$-0.200 < \Delta E < 0.125$	$-0.95 < \cos\theta_H < 1.00$	0.75
$\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$	$-0.200 < \Delta E < 0.150$	$-0.80 < \cos\theta_H < 1.00$	0.90
$\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$	$-0.200 < \Delta E < 0.150$	$-0.70 < \cos\theta_H < 1.00$	0.75

3.5.1 Multiple candidate treatment

It is possible, and indeed fairly common, that more than one candidate will pass all the selection criteria for a particular event, and it is therefore necessary to select one of the multiple candidates for fitting. On average, between 1.1 and 1.3 B candidates per event pass selection criteria in the on-peak i.e. signal region data-set. We select the best candidate by defining a χ^2 variable with inputs from the resonant masses and resolution of the η' candidates, or both the η' and the η candidates for $\eta'_{\eta\pi\pi}$ modes. The deviation of the candidate mass from the PDG value is divided by the resolution, and squared. For modes using both the η' and the η, the calculation is repeated for each mass and the sum of the individual χ^2 is used. The candidate with the lowest χ^2 is then selected. Any bias introduced by such a method is minimal, since only one variable here, $m_{\eta'}$, is fitted and thus a correlation is not possible. It can induce a small peak in the η' mass, but the background PDFs include this same shape. The values used to compute the χ^2 are shown in Table 3.3. The masses, m_0 are from the PDG [13] and the resolutions used are derived from data collected.

Table 3.3: Nominal values for resonance masses and resolutions used to select the best B candidate where multiple candidates exist.

Resonance	m_0 [GeV]	σ [GeV]
Modes with $\eta'_{\eta\pi\pi}$		
$\eta_{\gamma\gamma}$	0.54730	0.0113
$\eta'_{\eta\pi\pi}$	0.95778	0.0037
Modes with $\eta'_{\rho\gamma}$		
$\eta'_{\rho\gamma}$	0.95778	0.0073

3.5.2 An assessment of selection efficiency corrections

The selection efficiency is defined as the fraction of signal Monte Carlo events which pass the preselection criteria. However, modifications to this are required as there are known differences between the simulated and real data. These are summarised below:

- Tracking efficiency - efficiency modified down by 0.5-0.8%, with an associated systematic uncertainty of 1.3-1.4%, applied to all three daughter tracks of the resonances.
- Neutral correction - this corrects π^0 efficiency (-3%), and is also applied to modes containing an η (-2.7%) and each correction adds a systematic uncertainty of 3%.
- Single photon efficiency - this is applied to modes containing a $\rho\gamma$ as a result of the low momentum photon. This correction does not modify efficiency but adds a systematic uncertainty of 1.8%.
- K^0_s correction - the reconstruction of the K^0_s is adjusted by comparison with the average of all the K^0_s candidates which pass the selection criteria. A downward correction of 1.8% is applied to the subdecay channel $\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$, and a downward correction of 1.9% to $\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$. Systematic errors of the same magnitude are included in the tracking errors for each mode.

Chapter 4

The maximum likelihood fit model

An unbinned, extended, multi-variate maximum likelihood (ML) fit is at the heart of this analysis. It is performed after the skims and selection cuts described in Chapter 3. The cuts are loose to maximise efficiency and also to allow the background shapes to be well modelled with sufficient statistics using m_{ES} and ΔE sideband data. The success of the ML fit is dependent on the probability density functions (PDFs) for the physical observables being well determined for signal, $q\bar{q}$ continuum and $B\bar{B}$ backgrounds. The PDFs are assumed to be independent, but in practice correlations occur and must be handled with due care to ensure the bias on the fitted signal yields is quantified. The necessary fit validation required for this is studied in detail in Chapter 5.

4.1 Introduction

Consider an ensemble of N measurements of a set of observables x_i which are distributed according to a PDF $P(x;a)$, where a is the set of parameters of interest. The likelihood $\mathcal{L}$ of obtaining a particular data-set $x_1, x_2, ..., x_m$ for a particular value of parameters a, is given by the product of the individual likelihoods $P(x_i;a)$:

$$\mathcal{L} = \Pi P(x_i;a) \tag{4.1}$$

where the individual likelihoods are defined as:

$$P(x_i; a) = \Sigma_{j=1}^{k} n_j P_j(x_i) \tag{4.2}$$

where n_j is the number of events recorded in each hypothesis (signal, $B\bar{B}$ background or $q\bar{q}$ continuum background). $P_j(x_i)$ is the probability of an observable having a measured value x_i.

The object of the fit is to compute the value of a which maximises the probability of obtaining the observed data-set x_i. The quantities x_i are the observables ΔE, m_{ES}, $\mathcal{F}$, the masses of the resonances and the vector meson helicity angle, $\cos\theta_{\mathcal{H}}$. The quantities in a include the signal yield and the CP asymmetries, and the PDFs are constructed for signal, $q\bar{q}$ continuum and $B\bar{B}$ backgrounds.

Since $\mathcal{L}$ describes the relative probabilities of various values of x_i, maximising $\mathcal{L}$ allows the true value of the parameters of interest, a, to be determined. In practice, rather than maximising $\mathcal{L}$, it is more straightforward to minimise -ln $\mathcal{L}$ since the values of $\mathcal{L}$ are small.

The maximum likelihood can be computed as a set of k simultaneous equations, one per fitted parameter in a, with solutions:

$$\frac{d\mathcal{L}}{da_i} = 0, \quad i = 1, 2, .., k. \tag{4.3}$$

The statistical uncertainty and significance of the fit result are computed under the assumption that the maximum likelihood function for any fitted parameter a_i is Gaussian-like in the region close to the maxima, where $\mathcal{L}_{max} = \mathcal{L}(a_{true})$:

$$\mathcal{L}(a_i) = \mathcal{L}_{max} exp\left(-\frac{(a_i - a_{true})}{\sigma^2}\right). \tag{4.4}$$

The statistical uncertainty, σ, on the fitted parameters a is computed using this Gaussian-like assumption:

$$\frac{1}{\sigma^2} = \frac{d^2 ln\mathcal{L}}{da_i^2} \tag{4.5}$$

with the quoted value σ defined by the condition -2 ln $(\mathcal{L}(a_{true} + \sigma)/\mathcal{L}_{max}) = 1$. The significance of a fitted parameter is computed in the same manner, but with reference to the zero hypothesis:

$$\sigma^2 = -2\, ln\frac{\mathcal{L}_0}{\mathcal{L}_{max}}. \tag{4.6}$$

Normally, the integral of the likelihood function is normalised to unity. This is appropriate when a fixed number of events is expected. However, the likelihood fit used here must also account for the fact that the recorded number of events is subject to some random variation. This is achieved with an extended maximum likelihood fit, where the integral of the function is defined as the number of events which maximises the likelihood, rather than unity. The extended maximum likelihood is characterised by the use of an exponential factor to account for Poisson statistics in the (random) number of events recorded, N, and described as:

$$\mathcal{L} = \frac{e^{-\Sigma_j n_j}}{N!}\Pi_{i=1}^{N}\mathcal{L}_i \tag{4.7}$$

where i represents one of the N events, j is the number of hypotheses, represented here by signal, $B\bar{B}$ and continuum background components. The free parameter n_j represents the mean number of events in the sample for each hypothesis. $\mathcal{L}_{max}$ is the maximum value of the likelihood $\mathcal{L}$, which corresponds to the true value of a, and $\mathcal{L}_0$ is the value of $\mathcal{L}$ at $a = 0$.

4.2 Probability Density Functions (PDFs)

As mentioned above, the ML fit is dependent on an accurate characterisation of the signal and background components, continuum $q\bar{q}$ and charmless $B\bar{B}$. For each discriminating variable used, PDFs for signal and each of the backgrounds are constructed. For signal and charmless $B\bar{B}$ background fitting, simulated signal Monte Carlo (MC) data are used, and

for $q\bar{q}$ backgrounds, sideband data are used. An overview of the fit model, in particular the PDF shapes used to parameterise each fit component, are summarised in Table 4.1 and discussed further in the following Sections. Only the best candidate from each event is used for the purposes of constructing the PDFs, as discussed in Section 3.5.1. The PDFs for signal, background and continuum are presented in Section 4.4 for each of the two decay modes presented here, $\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$ and $\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$. The reader is referred to Section B.1 for the four other modes which contribute to the final combined result, $\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$, $\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$, $\eta'_{\eta\pi\pi}K^{*0}$ and $\eta'_{\rho\gamma}K^{*0}$. For information on parameters which are floated in the fit, see Section 4.3.2.

Variable	Signal	$q\bar{q}$ background	$B\bar{B}$ background
ΔE	Double Gaussian	P1 + P2	Gaussian + P1/P2
m_{ES}	Double Gaussian	ARGUS	ARGUS + Double Gaussian
$\mathcal{F}$	Bifur. Gaussian	Bifur. Gaussian + Gaussian	Bifur. Gaussian + Gaussian
$m_{\eta'}$	Double Gaussian	P2 + Double Gaussian	P2 + Double Gaussian
m_{K^*}	Breit-Wigner	P2 + Breit-Wigner	P2 + Breit-Wigner
$\cos\theta_{\mathcal{H}}$	P4	P4	Exponential + P2

Table 4.1: The PDF shapes used to fit the analysis variables described in Section 3.4, and the masses of the η' and K*. The shapes used comprise ARGUS functions, Gaussian and bifurcated Gaussians, Breit-Wigner functions, polynomials Px of order x, and an exponential function. These are discussed in Section 4.3.1.

Signal PDF samples

The signal PDFs are determined using simulated MC data. The resonances are fitted with either a Breit-Wigner function which describes their natural lineshapes, or if the natural width of the resonance is too narrow to be resolved, Gaussian distributions are used instead. Other quantities are fitted with appropriate functions; helicity angle with a polynomial and the Fisher discriminant with a bifurcated Gaussian.

Continuum background

As described in the previous Chapter, $q\bar{q}$ events ($u\bar{u}$, $d\bar{d}$, $s\bar{s}$ and $c\bar{c}$,) do not originate from a B meson and hence the reconstructed mass of these events show no peak at the B meson mass in the ΔE-m$_{ES}$ plane. An accurate characterisation of the background is obtained using m$_{ES}$ sideband data (m$_{ES} < 5.27$GeV) for all continuum PDFs with the exception of the m$_{ES}$ continuum PDF, where on-peak data in the ΔE sideband ($|\Delta E| > 0.1\,\mathrm{MeV}$) is used. The sidebands are used to generate the background PDFs to avoid the area where signal is expected. In the final fit the continuum background PDF also absorbs the similarly shaped $b \to c$ component of the $B\bar{B}$ background. Section 4.6 discusses this in detail.

$B\bar{B}$ background

Events which originate from a B meson decay, and which pass all selection cuts, but without being the signal modes of interest are called B background events. These mimic the final state of interest, and peak at the B meson mass. Many of these modes arise from the inclusion of a soft pion to a topologically-similar final state e.g. $B \to \eta' K$. The helicity angle cuts are a powerful tool in reducing the extent of these backgrounds, but it is not possible to reduce them to an insignificant level using this cut alone.

To account for these unwanted false signals, we split the recorded B background events into charmed ($b \to c$) and charmless components. The $b \to c$ component is absorbed into the continuum background PDFs described above, and so the $B\bar{B}$ PDFs are composed purely of the charmless component of the B background.

The charmless B background PDFs are constructed from MC simulation which is used to identify the relevant modes, and their significance to the final state of interest. A detailed discussion of the treatment of B backgrounds is found in Section 4.5.

4.3 Discriminating variable parameterisations

ΔE

A double Gaussian is used to parameterise the signal. A scale factor of 1.05 ± 0.05 is applied to the core Gaussian to correct the Monte Carlo data resolution, which is otherwise a good fit (no shift is applied). The shift and scale factor corrections used were confirmed by the study of a control sample of $B^- \to D^0\pi^-$, with $D^0 \to K\pi\pi^0$. This is done in order to explore the signal region using a topologically similar decay mode to the signal modes studies here, but without compromising the blind analysis approach. The discrepancies found between control sample MC and on-peak data are consistent with the corrections applied to the signal modes described here. The control sample study is described in Appendix A.

For continuum backgrounds we use m_{ES} sideband data and fit with a first or second order polynomial. The $B\bar{B}$ background is composed of a Gaussian, applied to any peaking component, and a polynomial of order one or two.

m_{ES}

A double Gaussian is used for the signal component of m_{ES}. The Monte Carlo data must be shifted to fit the peak seen in data as a result of small changes to the observed beam energy. The shift required varies with data-taking run, and is applied to both Gaussians as shown in Table 4.2 below. No scale factor is required. The error on the applied shift is the uncertainty in the CM energy due to beam energy spread.

The need for a shift due to the variable beam energy can be seen due to the dependence on it in the definition of m_{ES} in Equation 3.2 on page 61. As with ΔE, we confirm the magnitude of the m_{ES} shift is reasonable for the modes studied here via the $B^- \to D^0\pi^-$ control sample study detailed in Appendix A.

To fit the continuum background, we use on-peak data with a ΔE cut and fit this with an ARGUS function [35]. The $B\bar{B}$ background is fitted with the sum of an ARGUS and a double Gaussian for the peaking component. See Section 4.3.1 for a description of an

ARGUS function.

m_{ES}	Shift applied, MeV
Run 1	0.8 ± 0.2
Run 2	0.6 ± 0.2
Run 3	0.4 ± 0.2
Run 4	0.1 ± 0.2

Table 4.2: Shift applied to MC samples to correct for the differences between data and MC as a result of variability in the beam energy for each data taking run.

$\mathcal{F}$

The Fisher discriminant is fitted with the sum of a bifurcated Gaussian plus a standard Gaussian, included in order that the broad tails in the distribution are modelled accurately. m_{ES} sideband data are used to fit the continuum background. The background parameters are floated to allow the PDFs to adjust enough when moving to the full signal region, due to B background being absorbed. This allows the background parameters to be as accurate as possible.

Resonance masses

The signal PDFs for the resonance masses of the η' and K^* are obtained from signal MC samples. The signal PDF for the η' is fitted with a double Gaussian, the K^* with a Breit-Wigner function. The continuum and $B\bar{B}$ backgrounds are fitted with the sum of the actual signal PDF and a Chebychev polynomial, with the continuum background data being taken from the m_{ES} sidebands.

A scale factor of 1.14 and shift of 1.2 MeV is applied to correct the MC data for modes featuring the decay $\eta' \to \rho^0\gamma$. No such correction is required for the K^* or $\eta' \to \eta\pi\pi$.

K^* helicity angle: $\cos\theta_{\mathcal{H}}$

Both the signal and continuum background helicity angle PDFs are modelled with a fourth order polynomial. The continuum background PDF is derived from m_{ES} sideband data defined by $m_{ES} < 5.27\,\text{GeV}$ and is flatter than the signal PDF as a result of the presence of combinatoric resonances in addition to true resonances from generic production mechanisms.

The $B\bar{B}$ helicity angle PDF is fitted with a combination of an exponential function peaked at ± 1 and a second order polynomial. The exponential is required as a component of the $B\bar{B}$ backgrounds results from the inclusion of a soft pion to a hard kaon from the mode $\eta' K$. This shows strongly in the helicity angle PDF at values close to 1 since the kaon is "hard", travelling almost in line with the original B momentum. The daughter pion from the K^* is thus travelling against the B momentum.

4.3.1 Parameterisation shapes used

Gaussian, also bifurcated Gaussian

This is a function of the form:

$$f(x:\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) \tag{4.8}$$

which describes a variable x with width σ and mean μ. The variable x is often the measured mass, usually m_{ES}, or in the case of ΔE it is the missing event energy. The Gaussian is symmetric about the mean in standard form, however some distributions are fitted with a bifurcated Gaussian, which has a different widths either side of the mean, denoted σ_L and σ_R, where:

$$\sigma = \sigma_L \;;\; x < \mu \tag{4.9}$$

and

$$\sigma = \sigma_R \;;\; x > \mu\,. \tag{4.10}$$

ARGUS function

The ARGUS function [35] describes the shape of the m_{ES} continuum background PDF, and takes the form:

$$f(m_{ES} : \xi) = \frac{1}{N} m_{ES} \sqrt{1 - {m_{ES}}^2} \left[-\xi (1 - {m_{ES}}^2) \right] \tag{4.11}$$

where ξ is the slope of the ARGUS function, which is floated in the fit. The distribution is kinematically limited to half the centre of mass energy of 10.58 GeV. The ARGUS function assumes a uniform phase-space distribution of continuum background.

Breit-Wigner

This describes the natural lineshapes of resonances, and is employed where the natural width is large enough that the detector resolution is sufficient to reveal the natural lineshape. The function takes the form:

$$f(x : x_0 : \gamma) = \frac{1}{\pi} \left(\frac{\gamma}{(x - x_0)^2 + \gamma^2} \right) \tag{4.12}$$

where γ is the half-width at half-maximum, and x_0 is the central value of x. Where the detector resolution is not sufficient to reveal the natural lineshape of a resonance, fits using Gaussians are employed instead.

Polynomial of degree n

The polynomials used are taken from the orthogonal series of Chebychev expressions, chosen to minimise correlations, although care must be taken as they can yield negative values.

Parameter name	Description	Modes where param. is floated
deBkgP01	P1 coeff. for ΔE B background	All
deBkgP02	P2 coeff. for ΔE B background	All bar $K_s\pi^+$ modes
mesBkgc	ARGUS function slope, ξ of m_{ES}	All
fisBkgCasym	Asymmetry of bifur. Gaussian in $\mathcal{F}$	All
fisBkgCmean	Mean of bifur. Gaussian in $\mathcal{F}$	All
fisBkgCrms	Width of bifur. Gaussian in $\mathcal{F}$	All
helKstarBkgP01	P1 coeff. for $\cos\theta_{\mathcal{H}}$ background	$K_s\pi^0$ modes only
mEpBkgfracEp	η' param. weight in PDF	All
mEpPolyBkgP01	P1 coeff. for $m_{\eta'}$ background	All
mEpPolyBkgP02	P2 coeff. for $m_{\eta'}$ background	$K_s\pi^0$ modes only
mKstarBkgfracKst	K^* param. weight in PDF	All bar $K^+\pi^0$ modes
KstarMassBkgfracKs	K_s param. weight in PDF	$K^+\pi^0$ modes only
KstarMassPolyBkgP0	P0 coeff. for m_{K^*} background.	$K^+\pi^0$ modes only
mKstarPolyBkgP01	P1 coeff. for m_{K^*} background	All bar $K^+\pi^0$ modes
mKstarPolyBkgP02	P2 coeff. for m_{K^*} background	$K_s\pi^0$ modes only
nBkg	$q\bar{q}$ yield	All
nChmls	Charmless $B\bar{B}$ yield	All
nSig	Signal yield	All

Table 4.3: Floated parameters in the final fit. Note that not all parameters were floated for all modes. Where the floating of a parameter made the fit unstable, it was fixed in the fit.

4.3.2 Floating fit parameters

Where possible, the parameters in the maximum likelihood fit are allowed to float. Systematic uncertainties in the parameters are therefore included in the fit statistical error where possible. For example the background parameters can be determined from the full on-resonance data-set which is larger than the on-resonance sideband sample. Only parameters to which our signal yields are sensitive are floated, which includes the $q\bar{q}$ PDFs since these include the $b \to c$ background component, which increases when the full on-peak data

are used in place of the sideband sample. The initial values chosen for the floated parameters are the values determined by the PDFs.

The fitter was validated and shown to be able to converge with the number of degrees of freedom in the final fits by extensive use of toy MC validation (see Chapter 5). The parameters which were floated in the final fit are given below. Note that not all parameters were floated for all modes. The parameters floated for both toy studies and the final fit for each mode can be seen in Sections 6.2, 5.3 and B.7.

4.3.3 Correlations between discriminating variables

The basis of the maximum likelihood fit described in Section 4.1 is that the PDFs for all variables are uncorrelated. In practice this is not the case and correlations between the discriminating variables are examined by means of toy experiments, described in more detail in Chapter 5. The observed correlations are described in the correlation matrices and accounted for in the signal yields by correcting for the fit bias determined from the toy experiment studies, using embedded fully simulated Monte Carlo events.

4.4 PDF plots

The PDFs for the two principal modes follow. The PDFs for each of the fit variables is presented for signal MC, off-peak sidebands and $B\overline{B}$ MC.

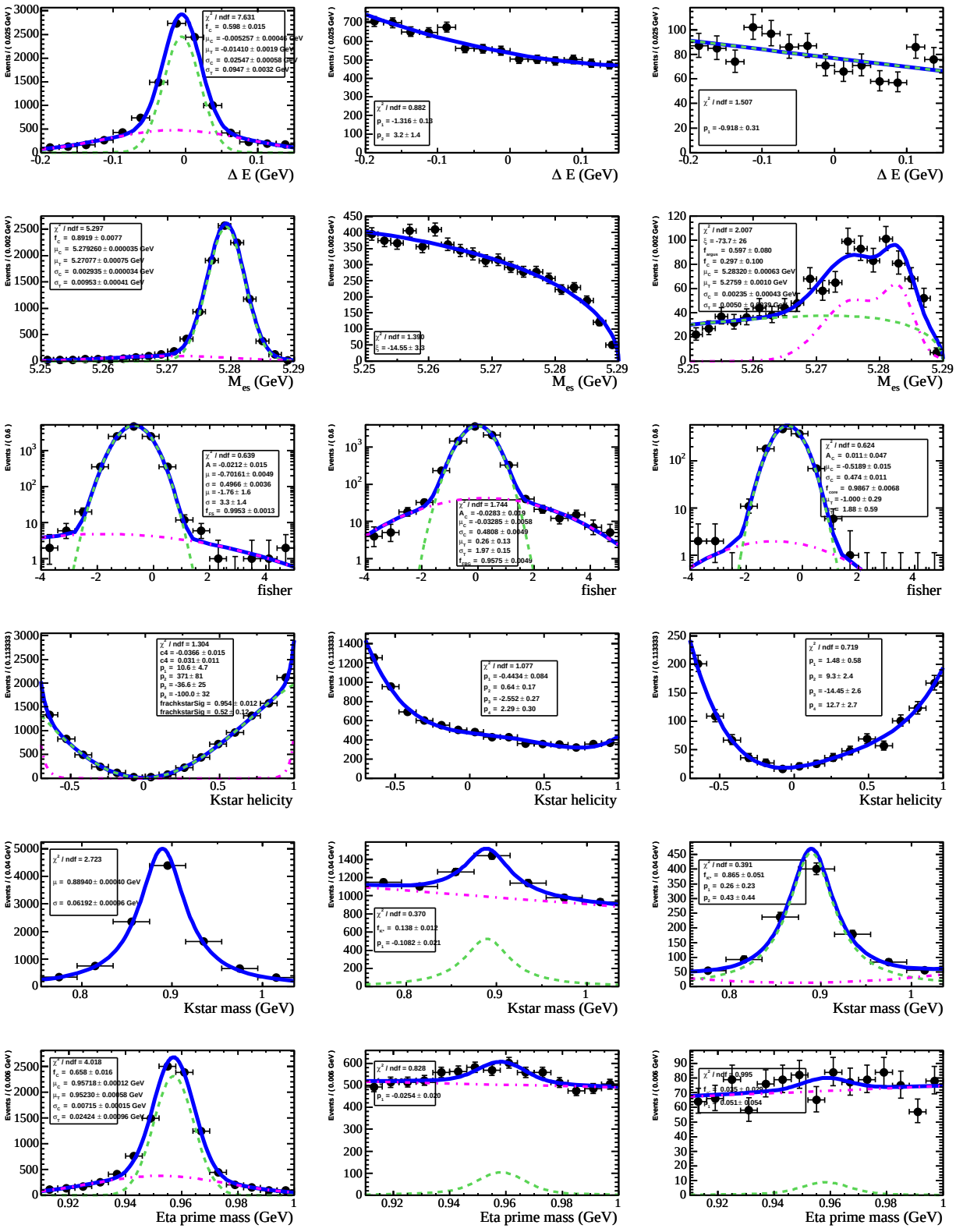

Figure 4.1: PDFs for the mode $\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$: Signal MC (*left*), on-peak sidebands (*centre*) and $B\bar{B}$ background (*right*). From top to bottom: ΔE, m_{ES}, $\mathcal{F}$, $\cos\theta_{\mathcal{H}}$, $m_{K^\star}$, $m_{\eta'}$. The solid blue curve is the sum of the dotted components described in Table 4.1.

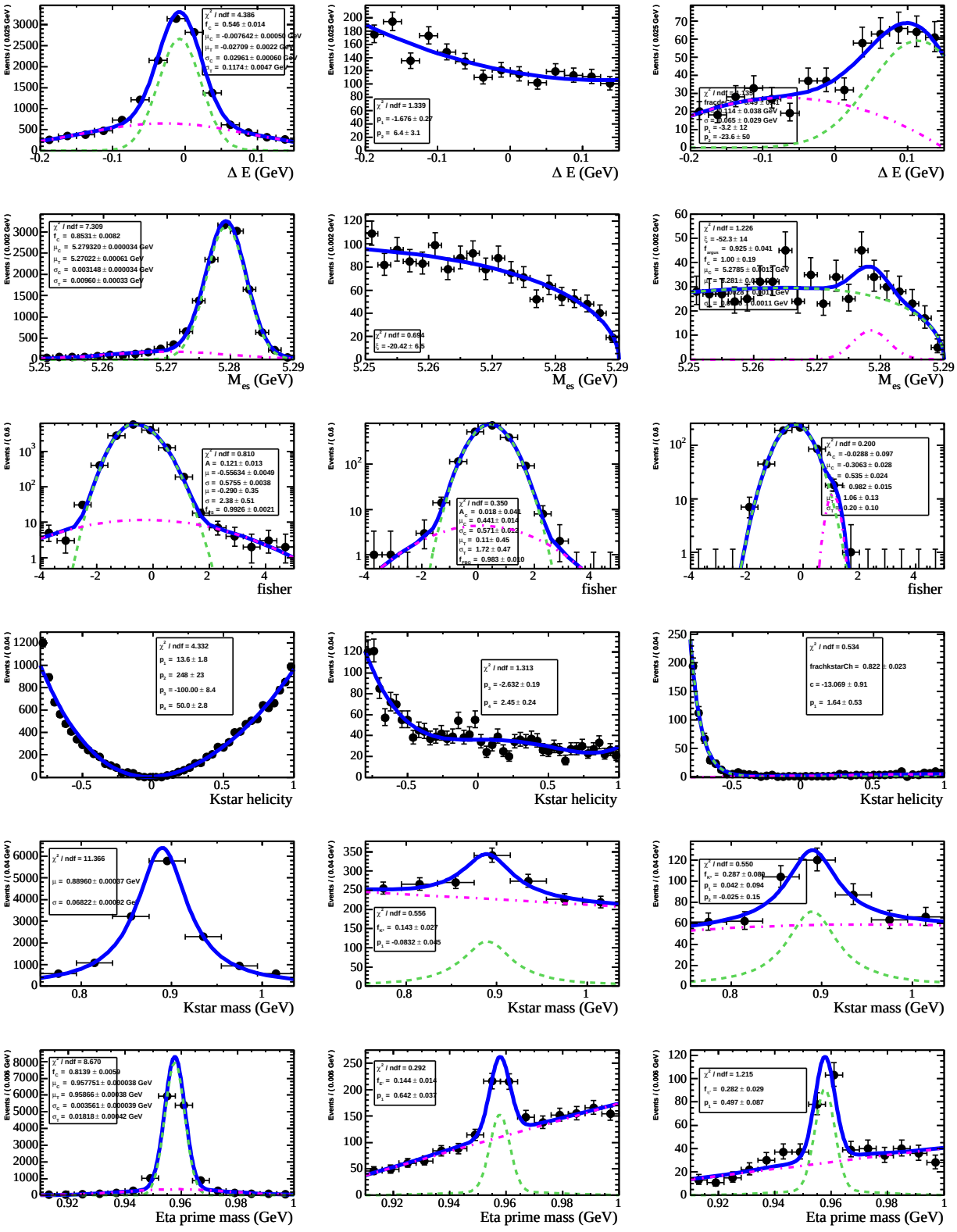

Figure 4.2: PDFs for the mode $\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$: Signal MC (*left*), on-peak sidebands (*centre*) and $B\bar{B}$ background (*right*). From top to bottom: ΔE, m_{ES}, $\mathcal{F}$, $\cos\theta_{\mathcal{H}}$, $m_{K^\star}$, $m_{\eta'}$. The solid blue curve is the sum of the dotted components described in Table 4.1.

4.5 $B\bar{B}$ background studies

4.5.1 Charmless $B\bar{B}$

Generic $B\bar{B}$ Monte Carlo simulation, containing all known B decay modes, was used to identify the charmless $B\bar{B}$ modes which passed the signal selection criteria. The events which passed the selection criteria were time-stamped, which was used to recover the event truth trees, to identify the mode in question. The frequency with which each decay mode in the generic $B\bar{B}$ passed the selection criteria was then used to produce a weighted "cocktail" $B\bar{B}$ data-set for each signal mode, from which the $B\bar{B}$ PDFs are derived.

Typically 125,000 Monte Carlo events of each charmless background were generated and mixed with appropriate weighting for the branching fraction and selection efficiency. Where available, Particle Data Group (PDG) [13] or HFAG [17] experimental measurements of branching fractions were used; where none were available estimates from similar decays were used instead, e.g. in the case of the decay $B \to a_1 K^*$, discussed below.

The final fitted charmless $B\bar{B}$ yields detailed in Table 6.1 (row labelled"BB yield") on page 103 are significantly higher than expected by the studies carried out here, shown in Tables 4.4 and 4.5 on the following pages for the two principal modes and in Section B.2 for the other four modes. Possible explanations for this are a leakage of charmed $B\bar{B}$ events into the charmless $B\bar{B}$ yield, and also evidence of this is found in Section 4.5.2. The leakage is not sufficient to explain the difference completely. Studies of non-resonant and higher resonance decays are considered and described below and ultimately a systematic error is assigned to cover the discrepancy.

Table 4.4: Charmless $B\overline{B}$ backgrounds to the $\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$ mode. For each background channel considered, we list the decay channel including all subdecays, production mode number, the selection efficiency (MC ϵ), an estimate for the branching fraction ($\mathcal{B}$), the daughter product branching fraction $\Pi\,\mathcal{B}_i$, the expected contribution of this channel in the on-peak data sample (Norm. #) and the number of simulated events used in the charmless cocktail (# in PDF).

Signal mode Bkg. channel	Mode #	MC ϵ (%)	Est. $\mathcal{B}$ (10^{-6})	$\prod\mathcal{B}_i$	Norm. # $B\overline{B}$ B kg.	# in PDF Bkg. file
$\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$						
$B^+ \to\ a_1^0K^{*+}(K^+\pi^0)(L, f_L = 1)$	5327	0.67	21	0.333	10.9	265
$B^0 \to\ a_1^-(\rho^0\pi^-)K^{*+}(K^+\pi^0)((L, f_L = 1)$	5325	0.66	42	0.167	10.8	263
$B^0 \to \rho^-K^{*+}_{K^+\pi^0}(T, f_L = 0.25)$	2500	0.93	9	0.333	6.4	157
$B^+ \to \rho^0K^{*+}_{K^+\pi^0}(L, f_L = 1)$	2355	0.59	10	0.333	4.6	111
$B^0 \to\ a_1^-(\rho^-\pi^0)K^{*+}(K^+\pi^0)(L, f_L = 1)$	5323	0.27	42	0.167	4.4	107
$B^0 \to\ a_1^0K^{*0}(L, f_L = 0.7)$	5329	0.05	21	0.467	1.1	28
$B^+ \to\ a_1^0\rho^+(L, f_L = 1)$	3999	0.01	48	1.000	1.1	26
$B^0 \to \rho^-K^{*+}_{K^+\pi^0}(L, f_L = 0.25)$	2499	0.46	3	0.333	1.1	26
$B^+ \to\ a_1^+(\rho^0\pi^+)K^{*0}(L, f_L = 0.7)$	5333	0.04	42	0.233	0.9	21
$B^0 \to\ a_1^+(\rho^+\pi^0)\rho^-(L, f_L = 1)$	4001	0.01	84	0.500	0.8	19
$B^+ \to\ a_1^+(\rho^+\pi^0)\rho^0(L, f_L = 1)$	4107	0.01	48	0.500	0.8	19
$B^0 \to\ a_1^+(\rho^0\pi^+)\rho^-(L, f_L = 1)$	4002	0.01	84	0.500	0.8	18
$B^+ \to \omega\ K^{*+}_{K^+\pi^0}(L, f_L = 1)$	2503	0.24	4	0.297	0.7	16
$B^+ \to\ a_1^+(\rho^+\pi^0)K^{*0}(L, f_L = 0.7)$	5331	0.03	42	0.233	0.6	14
$B^0 \to \eta'_{\rho\gamma}K^{*0}_{K^+\pi^-}$	2268	0.3	4	0.197	0.6	13
$B^0 \to\ a_1^0K^{*0}(T, f_L = 0.7)$	5330	0.03	21	0.200	0.3	3
$B^0 \to\ a_1^-(\rho^0\pi^-)K^+$	4871	0.01	13	0.667	0.3	3
$B^+ \to\ a_1^0K^+$	4874	0.02	6	1.000	0.3	3
$B^+ \to\ a_1^+(\rho^0\pi^+)K^{*0}(T, f_L = 0.7)$	5334	0.02	42	0.100	0.2	3
$B^+ \to\ a_1^+(\rho^0\pi^+)\rho^0(L, f_L = 1)$	4105	0	48	0.500	0.2	2
$B^0 \to \rho^0K^{*0}_{K^+\pi^-}(T, f_L = 0.5)$	2360	0.05	2.5	0.667	0.2	2
$B^+ \to \phi_{3\pi}K^{*+}_{K^+\pi^0}$	2294	0.01	10	0.333	0.1	1
$B^0 \to\ a_1^-(\rho^-\pi^0)K^+$	4960	0.01	13	0.333	0.1	1
$B^0 \to \rho^0K^{*0}_{K^+\pi^-}(L, f_L = 0.5)$	2359	0.03	2.5	0.667	0.1	1
$B^+ \to \rho^0\rho^+\pi^-$	4144	0	10	1.000	0.1	1
Total					47.5	1123

Table 4.5: Charmless $B\overline{B}$ backgrounds to the $\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$ mode. For each background channel considered, we list the decay channel including all subdecays, the production mode number, the selection efficiency (MC ϵ), a branching fraction estimate ($\mathcal{B}$), the daughter product branching fraction $\Pi\,\mathcal{B}_i$, the expected contribution of this channel in the on-peak data sample (Norm. #) and the number of simulated events used in the charmless cocktail (# in PDF).

Signal mode Bkg. channel	Mode #	MC ϵ (%)	Est. $\mathcal{B}$ (10^{-6})	$\prod \mathcal{B}_i$	Norm. # $B\overline{B}$ B kg.	# in PDF Bkg. file
$\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$						
$B^+ \to \eta'_{\eta_{\gamma\gamma}\pi\pi}K^+$	1506	0.17	69	0.174	4.8	342
$B^+ \to \eta_{3\pi}K^{*+}_{K^+\pi^0}$	1542	0.42	24	0.075	1.7	125
$B^0 \to \eta'_{\eta_{\gamma\gamma}\pi^+\pi^-}K^{*0}_{K^+\pi^-}$	2265	1.01	4	0.116	1.1	78
$B^0 \to\ a_1^-(\rho^-\pi^0)K^{*+}(K^+\pi^0)(L, f_L = 1)$	5323	0.02	42	0.167	0.3	18
Total					7.9	563

Estimation of branching fraction from $B \to a_1 K^*$ background

One major source of charmless B background is decays of the type $B \to a_1 K^*$. Since branching fractions for this decay had not yet been measured, an estimate of the expected contribution to our charmless background is used.

For the decays $B \to a_1 K^*$, we assume that

$$\frac{\mathcal{B}(B \to a_1^+ K^{*-})}{\mathcal{B}(B \to a_1^+ \rho^-)} = \frac{\mathcal{B}(B \to \rho^+ K^{*-})}{\mathcal{B}(B \to \rho^+ \rho^-)}. \tag{4.13}$$

We relate the branching ratio of the measured [36] decay channel $B^0 \to a_1\pi$ (33×10^{-6}) to the unmeasured $B^0 \to a_1\rho$ via QCD form factors for the π (131 MeV) and ρ (209 MeV). The relative magnitudes of the squared form factors allows an estimate of the branching ratio:

$$\mathcal{B}(B^0 \to a_1\rho) = \frac{209^2}{131^2}\mathcal{B}(B^0 \to a_1\pi) = 84 \times 10^{-6} \tag{4.14}$$

The RHS of Equation 4.13 is estimated to be around 1/3-1/2, from isospin relations which were used to estimate the branching ratio to the decay channel $B^0 \to \rho\rho$. The decay channel $B^0 \to K^*\rho$ [37] has an upper limit of $\mathcal{B}(B \to K^*\rho) < 12 \times 10^{-6}$, allowing an estimate of $B \to a_1 K^*$ to be quantified. 1/2 was chosen as the worst case, giving the largest backgrounds from $B \to a_1 K^*$.

It follows that

$$\mathcal{B}(B^0 \to a_1^+ K^{*-}) = 42 \times 10^{-6}. \tag{4.15}$$

We assign the same branching fraction to all modes with a charged $a_1^\pm$ in the final state.

For the modes with an a_1^0, we assume that they are pure penguins. In analogy to ρK^* decays, we use $\mathcal{B}(a_1^0 K^*)/\mathcal{B}(a_1^+ K^*) = \mathcal{B}(\rho^0 K^*)/\mathcal{B}(\rho^+ K^*) = 1/2$ and obtain for example

$$\mathcal{B}(B^+ \to a_1^0 K^{*+}) = 21 \times 10^{-6}. \tag{4.16}$$

We assume a polarisation of $f_L = 1.0$ for modes with a K^{*+} and $f_L = 0.7$ for modes with a $K^{\star 0}$.

4.5.2 $b \to c$ background studies

Rather than include a component in the fit which models charmed B decays explicitly, we allow the parameters of the $q\bar{q}$ background PDFs to float in the final fit so that any unmodelled B background is folded into the $q\bar{q}$ PDFs instead. This approach is taken since it is unlikely that many $b \to c$ decays will produce final states with signal-like topology due to the different quark content. However, it is necessary to explore the possibility that this approach introduces a fit bias on the signal yield by allowing $b \to c$ events to be included. This was done by examining the mode $\eta'_{\eta\pi\pi}\rho^+$, which is both topologically similar to the modes studied here, and the $\eta' \to \eta\pi\pi$ channel is particularly sensitive to $b \to c$ backgrounds. The results of the study are shown in Table 4.6.

A fit was run on an event sample comprising 0 signal events, 0 charmless $B\bar{B}$ events, 500 charmed $B\bar{B}$ events from generic $B\bar{B}$ Monte Carlo, with the remainder being drawn from the $q\bar{q}$ PDFs. Although the fitted charmless $B\bar{B}$ yields are affected by the charmed $B\bar{B}$ events, with around 10% of the embedded $b \to c$ events included in the charmless $B\bar{B}$ yields, the effect is not pronounced enough to completely explain the discrepancy in charmless $B\bar{B}$ yields described in Section 4.5.1. The signal yield is unaffected by the leakage of charmed $B\bar{B}$ events. We chose to embed 500 charmed $B\bar{B}$ events since our preselection criteria generated around that number of charmed events when run on generic B Monte Carlo.

Table 4.6: Summary of embedded toy experiments. The first row shows the result of toy experiments with the nominal fit for comparison. The second and third row show the result of toy experiments where we embed no signal events, 500 $b \to c$ events, and either 0 or 43 charmless events. The uncertainty in the fitted yields are determined by the RMS of the distribution divided by the $\sqrt{N_{experiments}}$. The uncertainties σ in the signal and backgrounds are the average uncertainties for the fitted signals yield.

Mode	N_{total}	N_{sig} (in)	$N_{B\bar{B}}$ (in)	$N_{b\to c}$ (in)	N_{sig} (fit)	$N_{B\bar{B}}$ (fit)	$\sigma(N_{sig})$ (fit)	$\sigma(N_{B\bar{B}})$ (fit)	bias [evts]
$\eta'_{\eta\pi\pi}\rho^+$	17287	0	43	0	-2.8 ± 0.7	50.7 ± 2.5	14.6	56	-2.8 ± 0.7
with 500 $b \to c$	17287	0	0	500	-2.8 ± 0.6	51.7 ± 2.4	12	50	-2.8 ± 0.6
with 500 $b \to c$	17287	0	43	500	-1.5 ± 0.8	100 ± 3	13	54	-1.5 ± 0.8

4.5.3 Backgrounds due to non-resonant $B \to \eta' K\pi$ decays and decays involving higher mass $K\pi$ resonances

A potential source of backgrounds are non-resonant B decays, and B decays involving higher mass $K\pi$ resonances. To examine the contribution from these decays, a study was conducted for some decay channels, from which a procedure was developed and applied to all other modes. MC for the decay $B \to \eta' K\pi$ was generated, with the $K\pi$ combination produced in S-wave since a phase-space decay was used. The non-resonant contributions are also included

in the tails of the distribution, away from the resonance. Background contributions from other non-S-wave sources, such as $K_1(1270)$, $K_1(1400)$, $K^*(1410)$ and $K_2^*(1430)$ are ignored because either the resonance is narrow and well separated from our signal region as is the case with the $K_2^*(1430)$, or because the daughter branching fraction to $K\pi$ is very small or zero, as with $K_1(1270)$, $K_1(1400)$ and $K^*(1410)$. The standard event selection criteria were applied, excluding the K^* mass and helicity angle requirements. The remaining events are reweighted and PDFs produced in the standard mass and helicity angle windows and a new component added to the nominal maximum likelihood fit. The results of the study are shown in Tables 4.7 and 4.8.

The nominal fit results are preserved when the S-wave component is fixed to zero. Floating the S-wave yield returns results which are consistent with zero. The signal yield is reduced in these fits, and the reduction is much more pronounced in the $\rho\gamma$ subdecay modes. Fixing the S-wave component to its fitted value does not change the reduced signal yields and restores the errors in the signal yields to close to their original relative amounts.

The pronounced drop in signal yield in the $\rho\gamma$ subdecay modes is due to a correlation between the S-wave and signal yields of around 30%, and between the S-wave and backgrounds of around 15%. Also, there may be some fitter-related effects as the S-wave yields in the $\rho\gamma$ modes have been observed to be unstable, varying depending on the starting value in the fit is changed.

Since the S-wave yields studied are consistent with zero, only a systematic error is included to account for the omission of a separate component in the final fit for all modes. Using the $\eta\pi\pi$ modes in combination, and efficiency-corrected S-wave yields, a branching fraction $\mathcal{B}(B \to \eta' K\pi) = 13^{+12}_{-10} \times 10^{-6}$ ($< 36 \times 10^{-6}$ @ 90% C.L.) is obtained.

	$\eta'_{\eta\pi\pi}K^{*0}$		$\eta'_{\rho\gamma}K^{*0}$	
	n_{Sig}	n_{swave}	n_{Sig}	n_{swave}
nominal fit	$22.6^{+7.7}_{-6.7}$	–	$35.1^{+14.2}_{-12.7}$	–
fix s-wave to 0	$21.7^{+7.6}_{-6.6}$	0	$34.3^{+14.1}_{-12.6}$	0
float s-wave	$19.8^{+10.4}_{-8.9}$	$6.2^{+10.3}_{-5.6}$	$16.5^{+24.9}_{-11.8}$	$69.8^{+12.4}_{-53}$
fix s-wave to fit	$19.8^{+7.5}_{-6.6}$	6.2	$16.3^{+12.9}_{-11.1}$	69.8
fix s-wave to expectation	$18.9^{+7.5}_{-4.4}$	9.4	$30.3^{+13.8}_{-12.4}$	13.5

Table 4.7: Results of including a $K^*(1430)$ in the fit to $\eta' K^{*0}$. The yields obtained are consistent with zero and hence only a systematic error is included in the final fit results to reflect any contribution from this potential background source.

	$\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$		$\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$	
	n_{Sig}	n_{swave}	n_{Sig}	n_{swave}
nominal fit	$11.2^{+5.7}_{-4.5}$	–	$14.8^{+11.2}_{-9.7}$	–
fix s-wave to 0	$11.2^{+5.7}_{-4.5}$	0	$14.9^{+11}_{-9.7}$	0
float s-wave	$9.8^{+6.2}_{-5.0}$	$4.8^{+10}_{-8.6}$	9.7^{+14*}_{-22*}	34.9^{+75*}_{-10*}
fix s-wave to fit	$9.8^{+5.6}_{-4.4}$	4.8	$7.9^{+11}_{-9.2}$	34.9
fix s-wave to expectation	$10.5^{+5.6}_{-4.4}$	2.3	$13.6^{+11.2}_{-9.6}$	5.6

Table 4.8: Results of including a $K^*(1430)$ in the fit to $\eta' K^{*+}(K^0_S\pi^+)$. Note that the fits for $\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$ with floating S-wave component do not have reliable error estimates (denoted by *).

Figure 4.3 shows the log-likelihood scan curves used to combine the two fit results and to derive the upper limit of $\mathcal{B}(B \to \eta' K\pi) < 36\times10^{-6}$. To determine the systematic uncertainty, we calculate the expected S-wave yield based on a central value of the branching fraction. We fix the S-wave yield to that value and run the fit. We take the systematic uncertainty as half the change in the signal yield between our nominal fit and the fit including the fixed S-wave yield.

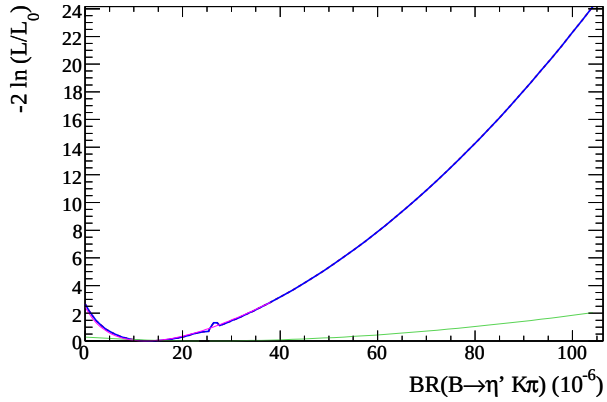

Figure 4.3: Negative log-likelihood (NLL) scan curves for the branching fraction of $B \to \eta' K\pi$ S-wave. The NLL curves only include the statistical uncertainty. The *thick blue* curve shows the NLL for the combination of the two measurements. The pink line is the scan curve for $\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$; the green line is the curve for $\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$.

Chapter 5

Fit validation

The fitting procedure involves many parameters, the floating of which can potentially make the fitter unstable. In addition, biases may be introduced by the fitting procedure itself, correlations between the variables used in the fit, and as a result of a limited statistical sample. Since the number of signal events is small compared to the overall data-set size, small biases can have a sizeable effect on the final signal yields. In addition, any bias in the fitter may also be a function of the signal yield. As such it is imperative to be satisfied that the fitter is indeed reliable, the magnitudes and stability of any fit biases are known and that the quoted uncertainties in fitted quantities can be relied upon, prior to performing the final fit to the on-peak data.

The first stage of the validation process was undertaken in two steps. Firstly, *pure* toys studies allowed any bias generated by the fitting procedure itself to be evaluated. Secondly, *embedded* toy studies allow the effect on the fit bias of correlations between the fitted variables themselves to be evaluated. The meaning of pure and embedded toy studies will now be explained.

5.1 Pure toy studies

All signal and background events generated for pure toy studies are derived from the fitted PDFs and hence take no account of any correlations between the variables. Pure toy studies are not affected by correlations between the input variables, allowing biases due to the fitting procedure itself to be quantified. Pure toy studies also provide an opportunity to check the fit robustness, in particular that the fit converged for each floated variable, with little observed bias. Failure of the fit to converge, or the production of large biases would indicate a problem with the fitting procedure, which in turn would jeopardise the reliability of any results produced by it. A fitting procedure which is shown to consistently converge with little bias in any of the floated variables will produce meaningful results.

Typically 1,000 toy experiments are run on each sample, with the samples containing the same number of signal, $q\bar{q}$ and $B\bar{B}$ events seen in the on-peak data samples in Table 6.1 on page 103. The signal yields and background parameters are allowed to float in the fit for each experiment. The distribution of fitted parameters from the experiments is examined for consistency. The pull, p, of the distribution is defined as:

$$p = \frac{x_{fit} - x_{expected}}{\sigma_{fit}} \tag{5.1}$$

where x_{fit} is the fitted value of a given parameter, $x_{expected}$ is the expected value of the parameter, and σ_{fit} is the fit error. The distribution of the pulls is expected to be Gaussian of zero mean and width unity. Should the distribution of pulls not conform to this, either the fit implementation is problematic, or that sources of error have not been quantified correctly.

For each sample of pure toy experiments, either zero or 10 signal events were included in the fit. The pure toy results for the two principal decay modes are shown in Table 5.1, and the pulls and parameter values are shown in Section 5.3. Results for the other four modes, used in the combined fit result, are shown in Sections B.6 and B.7.

It can be seen that when no signal events are generated, a small negative bias of a few events is observed, which is also observed in the embedded toy experiments described in the next

Section. The observed negative bias does not result from correlations between variables, since pure toy studies are insensitive to such correlations as the events are generated from the PDFs.

The bias is a feature of small signal yields and small sample sizes used in the maximum likelihood fit. Samples composed only of generated events from the background should have a fitted signal yield of zero. However, a small number of background events will be generated in the signal-like region. Since the probability of this is small, the expectation value of the number of such events, n_s is given by the binomial probability Np, where N is the number of events generated, and p is the probability of an event being signal-like. For samples where $n_s = Np$, the distribution can be fully described by the background PDF, and the signal yield should be near-zero.

However, the number of signal-like events in a sample is actually determined by the binomial distribution, and may be greater or less than Np:

$$P(n_s, N, p) = \frac{N!}{n_s!(N - n_s)!} p^{n_s} (1 - p)^{N - n_s}. \tag{5.2}$$

Where Np is small, the binomial probability of $n_s > Np$ is smaller than for $n_s < Np$, and the asymmetry results in a negative yield over a large number of experiments. When Np is larger, the probabilities become more symmetric as the distribution becomes more Gaussian-like than Poisson-like, and thus the bias is reduced. This effect has been studied extensively and is well understood.

5.2 Embedded toy studies

In contrast to the pure toy studies, embedded toy studies explore correlations between the variables of interest by using signal and $B\bar{B}$ events which are fully simulated events rather than generated from the PDFs. These correlations can affect the fitted signal yield by introducing a fit bias. Continuum background events were generated from the continuum PDFs in the same way as in pure toy studies, since it is not computationally feasible for

Mode	N_{total}	N_{sig} (in)	$N_{B\overline{B}}$ (in)	N_{sig} (fit)	$\sigma(N_{sig})$ (fit)	Bias [evts]
$\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$	3020	0	8	-4.6 ± 0.4	8.9	-4.6 ± 0.4
	3020	10	8	9.6 ± 0.3	5.9	-0.4 ± 0.3
$\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$	12996	0	48	-1.7 ± 0.5	10.7	-1.7 ± 0.5
	12996	10	48	9.9 ± 0.5	12.0	-0.1 ± 0.5

Table 5.1: Summary of results from pure toy MC studies for 0 and 10 signal events, and the expected number of charmless $B\overline{B}$ events included in each sample. In each case 200–1,000 toy experiments have been used, and the sample size for each mode, N_{total}, is chosen to match the size of the on-peak data-set for each mode, shown in Table 6.1 on page 103. The mean N_{sig} is taken from the average over all experiments, and its uncertainty is determined as the RMS of the distribution divided by $\sqrt{\text{N Experiments}}$. The value of $\sigma(N_{sig})$ is the average uncertainty for the fitted signal yield.

the full MC simulation data-set to contain enough continuum events to allow statistically-independent studies to be performed. A corresponding number of signal and charmless $B\bar{B}$ background events are sampled from the full MC simulation data-set and embedded into the data-sets to be fitted. All three components in the fit are allowed to float, and the observed fit bias is defined as the discrepancy between the expected fit yield, i.e. the number of embedded signal events, and the mean fitted signal yield.

Three samples were initially generated for each mode as part of the embedded toy studies, with each of these three samples containing the same number of events seen in the on-peak data samples in Table 6.1 on page 103. Up to 1,000 experiments were conducted on each of these three samples. The three samples for each mode were composed of differing quantities of signal and background events. One sample was generated with zero signal events, another with 10 signal events, and thirdly a set with 10 signal events but no charmless $B\bar{B}$ component. The results for the two principal decay modes are shown in Table 5.2. Results for the four other decay modes, used in the combined fit result, can be found in Section B.10 on page

156. The fourth sample, labelled "nFit-bias" was investigated after the fitter was used on the on-peak data sample.

Results from first sample, with zero signal events embedded and a $B\bar{B}$ component in the fit, shows the number of embedded $B\bar{B}$ events is consistent with the number of events in the on-peak data-set. The fitted signal yields show a small negative bias, as with pure toy studies.

The second sample includes the same $B\bar{B}$ number of events, and also includes 10 embedded signal events. The fitted signal yields are consistent with the number of embedded events, and the negative bias seen in the first sample has disappeared as expected.

Thirdly, the $B\bar{B}$ fit component was removed while retaining the same number of embedded signal and $B\bar{B}$ events. Some decay modes' signal yields showed little sensitivity to the removal of the $B\bar{B}$ component, e.g. $\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$. However other modes, such as $\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$, were found to have signal yields that are much more dependent. It was decided that the inclusion of a $B\bar{B}$ component in the fit was necessary for the decay $\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$.

The results from the first three samples are sufficient to demonstrate that the correlations between variables are well quantified, and modest in size. The benefit of the inclusion of a $B\bar{B}$ component in decay modes is also clear.

The final use of embedded toy studies occurred after the final fit to the full data-set, and the purpose of the studies was to quantify the fit bias systematic for the purposes of extracting the final branching fraction shown in Chapter 6. The final fitted signal yield for each mode for the full on-peak data-set is shown in Table 6.1 on page 103. These signal yields are reproduced in signal MC by an iterative process of embedded toy experiments. Quantities of bias-corrected signal events are embedded in order to match the recovered signal yield in the on-peak data-set as closely as possible. The difference between the number of embedded signal events and the number of fitted signal events from the real data-set is the fit bias, labelled "nFit-bias" in Tables 5.2 and B.10, which is used to calculate the branching fractions for the full results.

Table 5.2: Summary of embedded toy experiments. We include an additional line called "nFit-bias". This is a set of toys which give yields as close as possible to the yields found from the final fit. We use this for the final fit bias and systematic.

Mode	N_{total}	N_{sig} (in)	$N_{B\overline{B}}$ (in)	N_{sig} (fit)	$N_{B\overline{B}}$ (fit)	$\sigma(N_{sig})$ (fit)	$\sigma(N_{B\overline{B}})$ (fit)	bias [evts]
$\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$	3020	0	8	-4.5 ± 0.4	8.9 ± 0.4	9.3	16.5	-4.5 ± 0.4
	3020	10	8	10.8 ± 0.2	7.2 ± 0.7	4.6	16.4	$+0.8 \pm 0.2$
no $B\overline{B}$ fit	3020	10	8	10.8 ± 0.2	–	5.0	–	$+0.8 \pm 0.2$
nFit-bias	3020	4.5	46	5.5 ± 0.2	47 ± 0.8	4.9	17.2	$+1.0 \pm 0.2$
$\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$	12996	0	48	-3.2 ± 0.5	59.0 ± 2.4	11.2	53.6	-3.2 ± 0.5
	12996	10	48	10.8 ± 0.5	55.7 ± 3.0	11.9	65.2	$+0.8 \pm 0.5$
no $B\overline{B}$ fit	12996	10	48	11.7 ± 0.5	–	10.6	–	$+1.7 \pm 0.5$
nFit-bias	12996	5.5	337	3.2 ± 0.6	388.3 ± 3.0	11.7	66.5	-2.3 ± 0.6

5.3 Fit details

5.3.1 Correlations between variables

The correlations, ρ between the six variables in the fit are computed as follows:

$$\rho = \frac{V_{ij}}{\sigma_i \sigma_j} \tag{5.3}$$

where the covariance matrix, V_{ij} is defined as:

$$V_{ij} = \overline{x_i x_j} - \bar{x_i}\bar{x_j}. \tag{5.4}$$

The quantities $\sigma_{i,j}$ and $x_{i,j}$ are the standard deviation and central values of the parameters whose correlations are being evaluated. The covariance matrix contains 36 elements for this six-variable fit, but the six diagonal elements are trivially = 1, and the remaining 30 are duplicated which leaves 15 correlations to examine, presented in Tables 5.3 and 5.6, and are presented for signal MC, on-peak data and $B\overline{B}$ MC. Since the continuum events for

embedded toys were taken from the PDFs, as with pure toys, no correlations exist between the variables for the continuum PDFs. The parameter definitions are set out in Section 4.3.2 on page 76. It is clear that correlations between the variables used in the fit presented here, explored further with embedded toy experiments, are small or negligible across signal MC, on-peak data and $B\overline{B}$ MC. This confirms that the the use of a maximum likelihood fitting procedure, which relies on independent inputs, is justified.

Parameter values and parameter pulls obtained from pure toy studies described earlier are shown in Tables 5.4, 5.5, 5.7 and 5.8. These are by definition insensitive to correlations, but the distribution of pulls and the convergence of the fits demonstrate that the fitting procedure is robust.

$\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$

Correlation matrices, parameter values and parameter pulls

Table 5.3: Correlation matrices for signal MC, on-peak data and $B\bar{B}$ background MC for the mode $\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$.

	$\mathcal{F}$	m_{K^*}	$\cos\theta_{\mathcal{H}}$	ΔE	m_{ES}
Signal MC:					
m_{K^*}	-0.0036				
$\cos\theta_{\mathcal{H}}$	-0.0297	0.0240			
ΔE	-0.0215	0.0937	-0.0866		
m_{ES}	-0.0561	0.0034	0.1244	0.0228	
$m_{\eta'}$	-0.0033	0.0091	0.0006	-0.0094	-0.0032
On-Peak data:					
m_{K^*}	0.0110				
$\cos\theta_{\mathcal{H}}$	0.0635	0.0312			
ΔE	-0.0238	-0.0046	0.0235		
m_{ES}	0.0335	-0.0182	-0.0116	-0.0225	
$m_{\eta'}$	-0.0281	0.0190	-0.0337	0.0183	-0.0419
$B\bar{B}$MC:					
m_{K^*}	-0.0084				
$\cos\theta_{\mathcal{H}}$	0.0160	-0.0781			
ΔE	-0.0787	-0.0036	-0.1561		
m_{ES}	-0.0914	-0.0461	0.1747	0.0376	
$m_{\eta'}$	-0.0277	-0.0606	-0.0018	-0.0389	-0.0204

Table 5.4: Parameter values determined from pure toy studies for the mode $\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$. Where errors are signified as zero, they are < 0.005.

Parameter	mean	sigma
KstarMassBkgfracKs	0.14 ± 0.00	0.02 ± 0.00
KstarMassPolyBkgP0	-0.08 ± 0.00	0.04 ± 0.00
deBkgP01	-1.67 ± 0.01	0.21 ± 0.01
deBkgP02	6.41 ± 0.11	2.41 ± 0.08
fisBkgCasym	0.02 ± 0.00	0.03 ± 0.00
fisBkgCmean	0.44 ± 0.00	0.01 ± 0.00
fisBkgCrms	0.57 ± 0.00	0.01 ± 0.00
mEpBkgfracEp	0.14 ± 0.00	0.01 ± 0.00
mEpPolyBkgP01	0.64 ± 0.00	0.03 ± 0.00
mesBkgc	-20.51 ± 0.21	4.74 ± 0.15
nBkg	3007.08 ± 2.71	59.93 ± 2.01
nChmls	6.25 ± 0.73	16.30 ± 0.52
nSig	9.59 ± 0.26	5.87 ± 0.19

Table 5.5: Parameter pulls determined from pure toy studies for the mode $\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$.

Pulls	mean	σ_{fit}	mean err
KstarMassBkgfracKs	-0.04 ± 0.04	0.96 ± 0.03	0.0215
KstarMassPolyBkgP0	0.07 ± 0.05	1.01 ± 0.03	0.0351
deBkgP01	0.03 ± 0.04	0.97 ± 0.03	0.2161
deBkgP02	-0.06 ± 0.05	1.01 ± 0.03	2.3769
fisBkgCasym	-0.03 ± 0.04	0.99 ± 0.03	0.0301
fisBkgCmean	-0.03 ± 0.04	0.97 ± 0.03	0.0114
fisBkgCrms	-0.03 ± 0.05	1.06 ± 0.04	0.0082
mEpBkgfracEp	-0.01 ± 0.05	1.03 ± 0.03	0.0114
mEpPolyBkgP01	0.00 ± 0.05	1.05 ± 0.03	0.0296
mesBkgc	-0.02 ± 0.05	1.04 ± 0.03	4.5814
nBkg	0.08 ± 0.05	1.05 ± 0.04	57.2012
nChmls	-0.21 ± 0.05	1.06 ± 0.03	16.4276
nSig	-0.27 ± 0.06	1.23 ± 0.04	5.4821

$\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$

Correlation matrices, parameter values and parameter pulls

Table 5.6: Correlation matrices for signal MC, on-peak data and $B\bar{B}$ background MC for the mode $\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$.

	$\mathcal{F}$	m_{K^*}	$\cos\theta_{\mathcal{H}}$	ΔE	m_{ES}
Signal MC:					
m_{K^*}	-0.0116				
$\cos\theta_{\mathcal{H}}$	-0.0224	0.0165			
ΔE	-0.0028	0.1062	-0.0977		
m_{ES}	-0.0491	0.0150	0.0784	-0.0042	
$\mathrm{m}_{\eta'}$	0.0093	-0.0015	0.0034	0.0152	-0.0005
On-Peak data:					
m_{K^*}	-0.0326				
$\cos\theta_{\mathcal{H}}$	0.0278	0.0332			
ΔE	0.0043	0.0052	0.0138		
m_{ES}	-0.0062	0.0080	0.0223	-0.0003	
$\mathrm{m}_{\eta'}$	0.0045	-0.0110	-0.0019	0.0020	0.0207
$B\bar{B}$MC:					
m_{K^*}	0.0436				
$\cos\theta_{\mathcal{H}}$	-0.0527	0.0357			
ΔE	0.0373	-0.0222	-0.0684		
m_{ES}	-0.0287	-0.0137	0.0552	0.0102	
$\mathrm{m}_{\eta'}$	-0.0528	0.0332	-0.0250	-0.0244	0.0321

Table 5.7: Parameter values determined from pure toy studies for the mode $\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$. Where errors are signified as zero, they are < 0.005.

Parameter	mean	sigma
KstarMassBkgfracKs	0.14 ± 0.00	0.01 ± 0.00
KstarMassPolyBkgP0	-0.11 ± 0.00	0.02 ± 0.00
deBkgP01	-1.31 ± 0.00	0.09 ± 0.00
deBkgP02	3.19 ± 0.05	1.08 ± 0.03
fisBkgCasym	-0.03 ± 0.00	0.02 ± 0.00
fisBkgCmean	-0.03 ± 0.00	0.01 ± 0.00
fisBkgCrms	0.48 ± 0.00	0.00 ± 0.00
mEpBkgfracEp	0.05 ± 0.00	0.01 ± 0.00
mEpPolyBkgP01	-0.02 ± 0.00	0.02 ± 0.00
mesBkgc	-14.68 ± 0.10	2.21 ± 0.07
nBkg	12943.49 ± 5.68	126.37 ± 4.13
nChmls	47.79 ± 2.94	65.56 ± 2.11
nSig	9.88 ± 0.54	11.96 ± 0.39

Table 5.8: Parameter pulls determined from pure toy studies for the mode $\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$.

Pulls	mean	σ_{fit}	mean err
KstarMassBkgfracKs	0.03 ± 0.04	1.00 ± 0.03	0.0102
KstarMassPolyBkgP0	-0.05 ± 0.04	0.97 ± 0.03	0.0168
deBkgP01	0.06 ± 0.04	0.93 ± 0.03	0.0995
deBkgP02	-0.03 ± 0.05	1.01 ± 0.03	1.0663
fisBkgCasym	0.05 ± 0.04	1.00 ± 0.03	0.0154
fisBkgCmean	0.02 ± 0.05	1.04 ± 0.03	0.0051
fisBkgCrms	-0.03 ± 0.05	1.02 ± 0.03	0.0035
mEpBkgfracEp	0.01 ± 0.05	1.04 ± 0.04	0.0085
mEpPolyBkgP01	0.02 ± 0.05	1.01 ± 0.03	0.0158
mesBkgc	-0.06 ± 0.04	0.96 ± 0.03	2.3014
nBkg	0.04 ± 0.04	0.98 ± 0.03	128.5505
nChmls	-0.04 ± 0.05	1.04 ± 0.03	63.2639
nSig	-0.14 ± 0.05	1.22 ± 0.04	10.4745

Chapter 6

Results

6.1 Branching ratio results

The results of the maximum likelihood fit, along with a summary of the analysis on the "unblinded" 210.5 fb^{-1} on-peak data-set can be seen in Table 6.1 on the following page. The fit was performed using the PDF shape parameters determined from signal and $B\overline{B}$ MC studies discussed in Chapters 4 and 5, although the $q\overline{q}$ PDF parameters were floated to allow any $b \to c$ background to be folded into this component. The results for the two modes studied here are highlighted below, and then the combined results are given. All upper limits are expressed at 90% confidence:

$$\begin{aligned} \mathcal{B}(B^+ \to \eta'_{\eta\pi\pi} K^{*+}_{K^+\pi^0}) &< 9.5 \times 10^{-6} \\ \mathcal{B}(B^+ \to \eta'_{\rho\gamma} K^{*+}_{K^+\pi^0}) &< 22.2 \times 10^{-6}. \end{aligned}$$

The four-mode combined charged channel result is:

$$\mathcal{B}(B^+ \to \eta' K^{*+}) < 7.9 \times 10^{-6}$$

and the final six-mode combined fit is found to be:

$$\mathcal{B}(B \to \eta' K^*) \;=\; (4.1 \pm 1.0 \pm 0.5) \times 10^{-6}.$$

The combined charged mode $\eta' K^{*+}$ has a significance of 3.6 σ, and the total combined K^* fit has a signficance of 5.6 standard deviations. The log-likelihood scan curves for each mode, and the combined modes are shown in Figures 6.1 and 6.2 in Section 6.3. The conversion from fitted numbers of events to a branching ratio is explained in Section 6.2.

ML fit quantity	$\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$	$\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$	$\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$	$\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$	$\eta'_{\eta\pi\pi}K^{*0}$	$\eta'_{\rho\gamma}K^{*0}$
$\cos\theta_{\rm T}$ cut	0.9	0.75	0.9	0.75	0.9	0.75
Data combs	1.18	1.13	1.29	1.21	1.13	1.09
Signal MC combs	1.31	1.13	1.42	1.21	1.22	1.14
MC self XF frac.	0.23	0.22	0.36	0.28	0.19	0.19
Events to fit						
On-peak	2114	9962	3020	12996	4837	23790
Signal yield	$11.2^{+5.7}_{-4.5}$	$14.8^{+11.2}_{-9.7}$	$5.2^{+5.4}_{-3.6}$	$3.1^{+12.1}_{-9.6}$	$22.6^{+7.7}_{-6.7}$	$35.1^{+14.2}_{-12.7}$
BB yield	$16.5^{+11.1}_{-9.4}$	228^{+72}_{-71}	46^{+22}_{-21}	337^{+74}_{-72}	$45.8^{+24.3}_{-22.3}$	396^{+115}_{-115}
ML-fit bias (events)	+0.8	+2.9	+1.0	−2.3	+1.7	+9.5
MC ϵ (%)	19.2	16.4	11.6	8.4	20.1	17.2
Tracking corr. (%)	97.9	98.5	97.9	98.5	97.4	98.0
K^0_S corr. (%)	98.2	98.1	–	–	–	–
Neutrals corr. (%)	97.3	100.0	94.3	97.0	97.3	100.0
Corr. ϵ (%)	18.0	15.8	10.7	8.0	19.0	16.9
$\prod \mathcal{B}_i$ (%)	4.0	6.8	5.8	9.8	11.6	19.7
Corr. $\epsilon \times \prod \mathcal{B}_i$ (%)	0.7	1.1	0.6	0.8	2.2	3.3
Stat. sign. (σ)	3.3	1.3	1.2	0.5	4.1	2.2
Signif. w syst. (σ)	3.2	1.2	1.2	0.5	3.9	2.0
$\mathcal{B}(10^{-6})$	$6.2^{+3.4}_{-2.7}$	$4.7^{+4.5}_{-3.9}$	$2.9^{+3.7}_{-2.6}$	$2.9^{+6.7}_{-5.4}$	$4.1^{+1.5}_{-1.3}$	$3.3^{+1.9}_{-1.6}$
UL $\mathcal{B}$ (10^{-6})	11.6	11.9	9.5	22.2	6.4	6.4
Combined results						
$\mathcal{B}(10^{-6})$		$4.9^{+1.9}_{-1.7} \pm 0.8$			$3.8 \pm 1.1 \pm 0.5$	
Stat. sign. (σ)		3.7			4.7	
Signif. w syst. (σ)		3.6			4.3	
UL $\mathcal{B}(10^{-6})$		7.9			–	
$\mathcal{A}_{ch}$		$0.30^{+0.33}_{-0.37} \pm 0.02$			$-0.08 \pm 0.25 \pm 0.02$	

Table 6.1: Branching fraction results for all six modes used in the fit for $B \to \eta' K^*$.

With reference to Table 6.1 on the previous page, the quantities shown are:

$\cos\theta_{\mathrm{T}}$ cut

A tighter $\cos\theta_{\mathrm{T}}$ cut is used in modes where the decay $\eta' \to \rho\gamma$ occurs, since they are more susceptible to the jet-like backgrounds discussed in Section 4.2. The cuts applied are displayed in Table 3.2 on page 65.

Data combs

This quantifies the average number of B candidates which pass the selection criteria per event. Only one candidate per event is retained, and the process for choosing the best candidate is described in Section 3.5.1 on page 65.

MC self cross-feed fraction

This was computed by comparing the total number of selected signal events to the number of truth matched events, using signal Monte Carlo. It is a measure of the fraction of selected signal events which have not been matched to appropriate MC truth candidates.

Number of events to fit

This is the number of events which pass the skims and pre-selection/event shape cuts described in Chapter 3, the proportion of which is shown in row 9 as MC ϵ (%). These events are used to make the PDFs shown in Section 4.2.

Signal/BB yield and bias

These yields are derived from the maximum likelihood fit outlined in Chapter 4, and the bias shown in row 8 is derived from the embedded toy studies described in Section 5.2.

Tracking, neutral particle and K_s efficiency

The MC ϵ (%) is combined with other MC efficiency corrections described in Section 3.5.2. For each of the three tracks(two daughter pions and the K^*), a tracking efficiency correction of either $-0.5\,\%$ or $-0.8\,\%$ is applied.

For neutral particles, the correction factor is a linear sum of $-2.7\,\%$ for subdecays containing an $\eta\pi\pi$ and $-3.0\,\%$ for modes containing a π^0. Single photons in the decay $\eta \to \rho\gamma$ incur a systematic error but no efficiency correction.

Correction factors are applied to K_s reconstruction to account for differences between data and MC. The correction required varies by data-taking run and by decay channel, and is luminosity-weighted when combined. A systematic is assigned to the two modes which feature a K_s to reflect the required corrections.

Product branching fraction $\prod \mathcal{B}_i$ corrections

The product branching fraction must be included since the MC is constrained to simulate only the decay chain of interest.

Statistical significance, σ and upper limits, UL

This is calculated from the branching ratio fit, and is shown both prior to and post the inclusion of systematic errors. The upper limits are are 90% confidence limits for each decay mode, which are then combined and displayed in the row below.

6.2 Maximum likelihood fit results

The fit results for the two principal modes are presented in Tables 6.2 and 6.3 in the following pages. The fit results for the other four modes used in the combined fit result can be found in Section B.10 on page 156.

The branching fractions for each decay mode is calulated as follows:

$$\mathcal{B} = \frac{Y - Y_b}{\epsilon \, N_B \, \Pi \, \mathcal{B}_i} \tag{6.1}$$

where Y is the fitted signal yield, Y_b is the fit bias, ϵ is the MC efficiency, N_B is the number of B mesons of the correct charge produced, and $\Pi\,\mathcal{B}$ is the normalised daughter product branching fraction, taken from the PDG.

Table 6.2: Maximum likelihood fit results for $\eta'_{\eta\pi\pi} K^{*+}_{K^+\pi^0}$. The parameter GblCorr is the global correlation coefficient, defined for a particular variable as the linear combination of correlations which maximises the correlation between that variable and all others.

Floating Parameter	InitialValue	FinalValue	(+HiError,-LoError)	GblCorr.
KstarMassBkgfracKst	0.14278	0.11953	$(+0.0219, -0.0217)$	0.059054
KstarMassPolyBkgP01	-0.083234	-0.095821	$(+0.0350, -0.0349)$	0.035636
deBkgP01	-1.6762	-2.0299	$(+0.219, -0.217)$	0.235496
deBkgP02	6.4442	5.8938	$(+2.50, -2.37)$	0.169477
fisBkgCasym	0.018479	0.067513	$(+0.0326, -0.0325)$	0.155374
fisBkgCmean	0.44084	0.46667	$(+0.0119, -0.0118)$	0.383147
fisBkgCrms	0.57121	0.56891	$(+0.00852, -0.00838)$	0.260035
mEpBkgfracEp	0.14377	0.13983	$(+0.0116, -0.0114)$	0.098723
mEpPolyBkgP01	0.64207	0.60543	$(+0.0303, -0.0310)$	0.025999
mesBkgc	-20.416	-17.384	$(+4.65, -4.65)$	0.070830
nBkg	3,541	2,968.7	$(+58.6, -57.9)$	0.331051
nChmls	8	46.094	$(+22.3, -21.0)$	0.535634
nSig	10	5.2080	$(+5.40, -3.69)$	0.187360

Table 6.3: Maximum likelihood fit results for $\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$. The parameter GblCorr is the global correlation coefficient, defined for a particular variable as the linear combination of correlations which maximises the correlation between that variable and all others.

Floating Parameter	InitialValue	FinalValue	(+HiError,-LoError)	GblCorr.
KstarMassBkgfracKst	0.13763	0.11770	$(+0.0105, -0.0105)$	0.322742
KstarMassPolyBkgP01	-0.10821	-0.099477	$(+0.0168, -0.0169)$	0.052856
deBkgP01	-1.3155	-1.2861	$(+0.102, -0.101)$	0.281371
deBkgP02	3.1954	3.0721	$(+1.10, -1.08)$	0.286082
fisBkgCasym	-0.028291	-0.024284	$(+0.0154, -0.0155)$	0.247535
fisBkgCmean	-0.032853	-0.021794	$(+0.00530, -0.00531)$	0.494438
fisBkgCrms	0.48080	0.47528	$(+0.00373, -0.00372)$	0.304830
mEpBkgfracEp	0.050639	0.070342	$(+0.00879, -0.00874)$	0.094932
mEpPolyBkgP01	-0.020749	-0.0017559	$(+0.0164, -0.0164)$	0.063185
mesBkgc	-14.552	-13.924	$(+2.35, -2.36)$	0.272549
nBkg	$3,541$	$12,655$	$(+132, -130)$	0.487754
nChmls	337.00	336.97	$(+73.7, -72.2)$	0.713680
nSig	6.0000	3.1422	$(+12.1, -9.64)$	0.373077

6.3 Log-likelihood scan curves

The likelihood scan curves for each mode are presented in Figure 6.1, and the combined plots in Figure 6.2. The curves for each mode are shown before and after the inclusion of systematic errors.

There is a distinctive gradient-discontinuity in the likelihood function curves for the modes $\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$ and $\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$. This is a result of floating the charmless yield. The charmless yield rises significantly around the region of the discontinuity, albeit still within the quoted uncertainties. The fitter converges on a second minimum in this region. This does not affect the overall fit results and the quoted errors include the uncertainties as a result of this effect.

The next step involves the production of combined scan curves for the four-mode overall $\eta' K^{*+}$ result, and the six mode combined charged and neutral $\eta' K^{*0}$ branching fraction result. When combining likelihoods, care must be exercised in the treatment of correlated and uncorrelated systematic errors detailed in Chapter 7. The scan curves without the inclusion of systematic errors already reflect statistical uncertainties in the number of events, as discussed in Chapter 5.

The combined statistical error for a decay channel is naturally obtained by combining the log-likelihoods of the subdecay modes. However, the combined systematic errors require more careful treatment since some of these errors are correlated between several subdecay modes, and others are uncorrelated. To evaluate the combined systematic errors correctly, each subdecay mode's likelihood is convolved with a Gaussian function representing the uncorrelated systematic errors for that decay channel. Once the uncorrelated errors for each subdecay mode have been included in the log-likelihoods, the subdecay channels can be combined. After the combination of the subdecays, the correlated systematic errors can be applied to the combined likelihood, since these affect subdecays in the same manner. The resultant likelihood scan curve contains the total statistical and uncorrelated and correlated systematic errors.

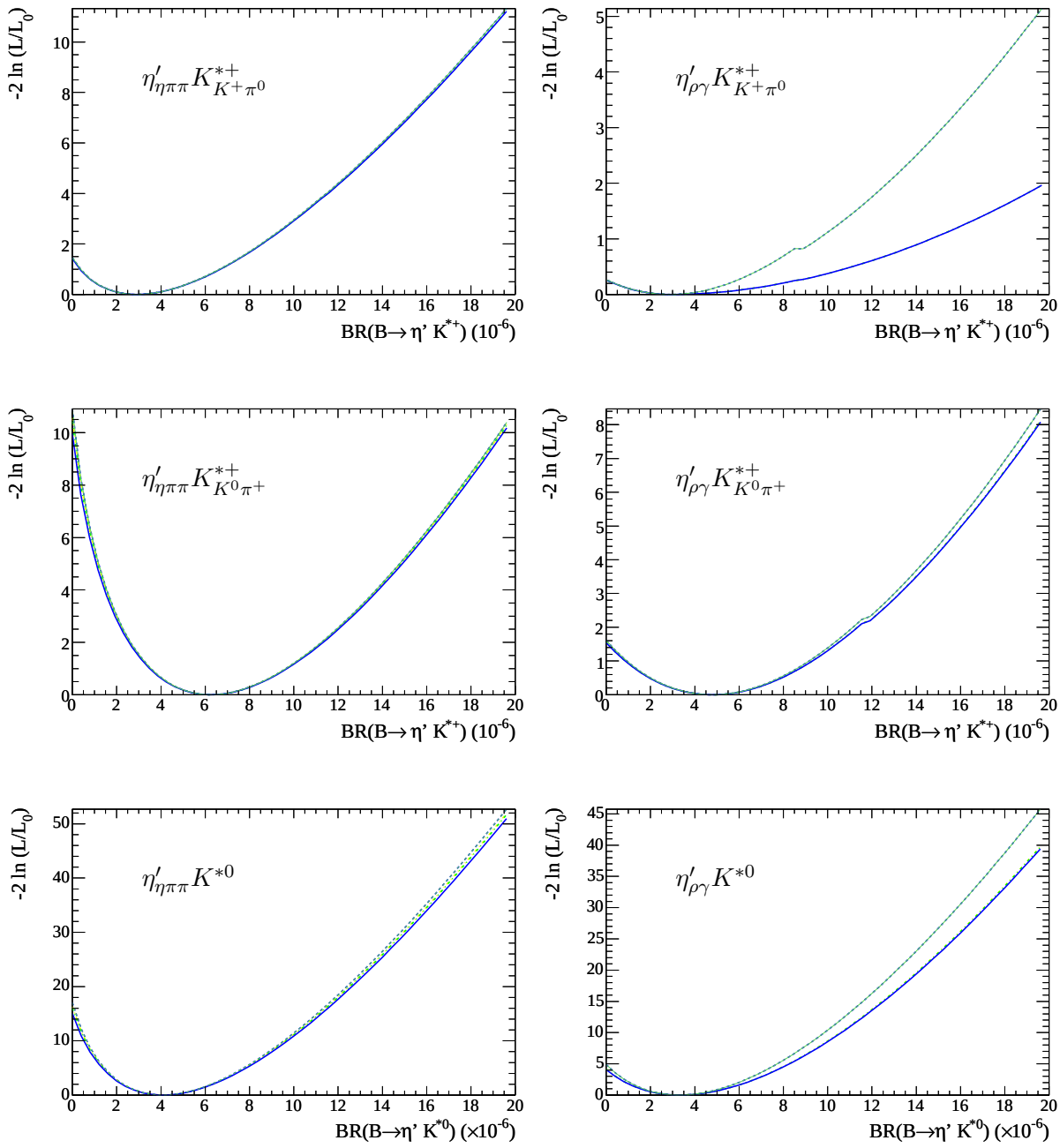

Figure 6.1: Negative log-likelihood (NLL) scan curves for all modes. The bold blue curve (lowest curve in each plot) shows the curve after all systematic errors have been taken into account, whereas the green curve contains only the statistical error. The significance (expressed in standard deviations, σ) of the result for each mode is computed as the square root of the curve's intersection on the y-axis. The gradient discontinuity seen in the scan curves for $\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$ and $\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$ is related to the floating of the charmless yield.

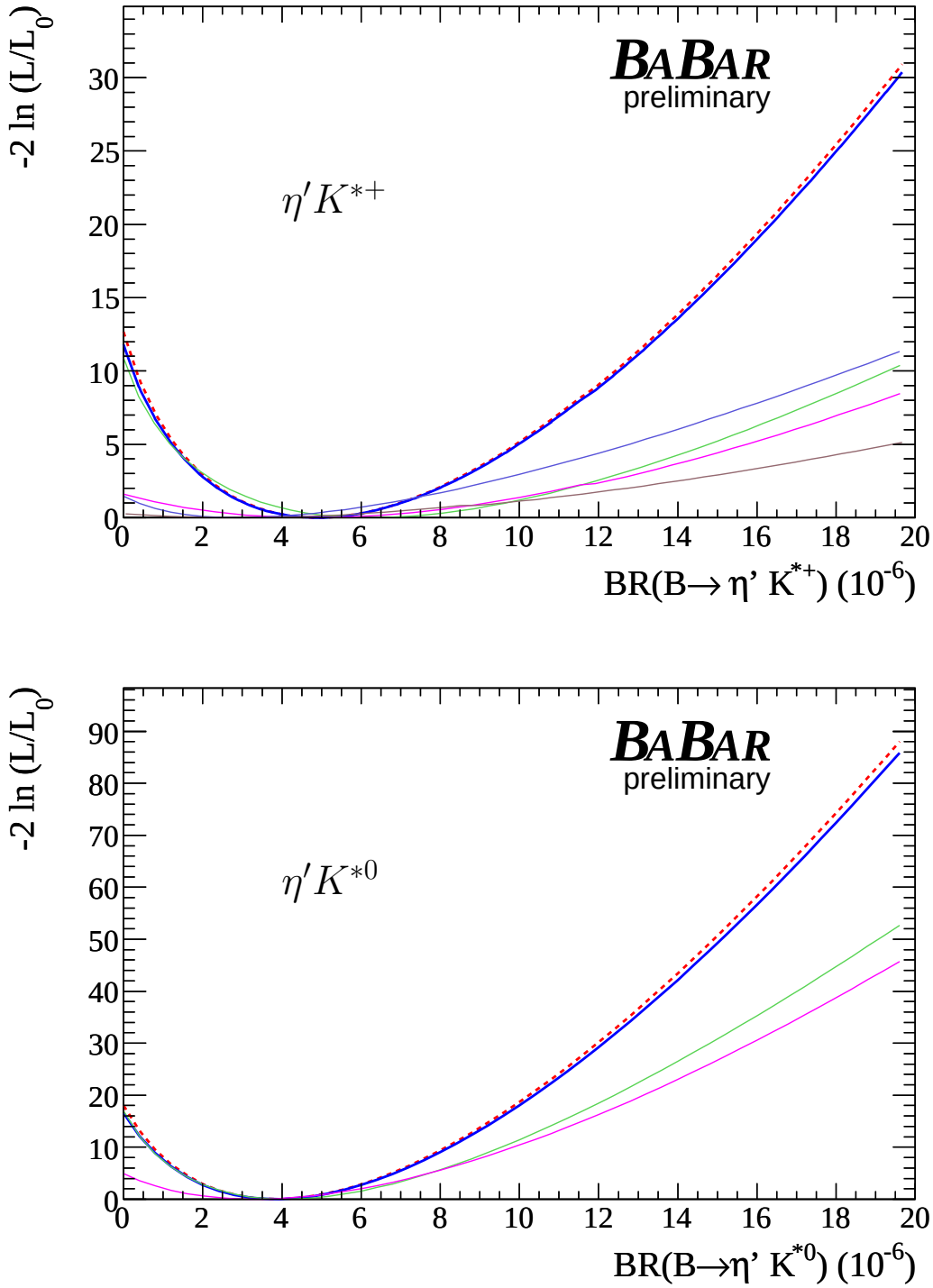

Figure 6.2: Negative log-likelihood scan curves for the combined modes. Individual sub-decay modes are the thin coloured lines. In the top plot these are: blue = $\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$; green = $\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$; pink = $\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$; grey = $\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$. In the bottom plot these are: pink = $\eta'_{\rho\gamma}K^{*0}$; green = $\eta'_{\eta\pi\pi}K^{*0}$. The solid blue curve is the final result and the dotted lines are without correlated and uncorrelated systematic errors.

6.4 Cross-checks

6.4.1 Likelihood ratio distributions

The likelihood ratio $\mathcal{L}_{sig}/[\mathcal{L}_{sig} + \sum \mathcal{L}_{bkg}]$ for both data and Monte Carlo is presented for the two principal decay channels in Figure 6.3. The plots for the other four modes, used for the combined fit, can be found in Section B.5. The points represent the on-resonance data, the red histograms are derived from toy MC generated from the PDFs of both background components and the green histograms are toy MC generated from the signal PDFs. Since the signal yields are small, the green histograms are barely visible and are unsurprisingly found at the unity side of the ratio plots. There is good agreement between the on-resonance data points and the summed histograms, which demonstrates that the likelihood ratio distribution is well described by the fitter. The total number of on-peak events in each histogram is found in Table 6.1 in the line "Events to fit".

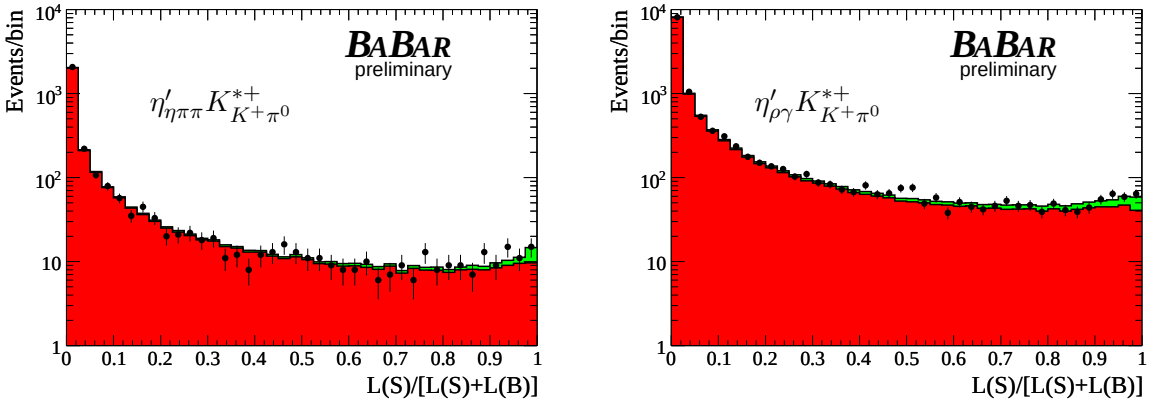

Figure 6.3: Plots of the likelihood ratio $\mathcal{L}_{sig}/[\mathcal{L}_{sig} + \sum \mathcal{L}_{bkg}]$ for all modes. The points are the on-resonance data, the red histograms are background toy ($q\overline{q}$ & $B\overline{B}$) and the green histograms are signal toy. The on-resonance points are in good agreement with the toy histograms.

6.4.2 Projection plots

The following pages display projection plots of the data over each of the fit variables, onto which the result of the combined fit has been overlaid for the two principal modes. Projection plots for the other four modes used in the combined fit can be found in Section B.4 on page 142. These plots are obtained using a likelihood ratio $\mathcal{L}_{sig}/[\mathcal{L}_{sig} + \sum \mathcal{L}_{bkg}]$ cut of 0.9 on the data. Figures 6.3 and B.5 show that this cut will preserve most of the signal events, shown in green, while discarding a large fraction of the red background events. The resulting data-set is a signal-enhanced subset of the full data-set, although since the signal yields are very small, the projection plots are still comprised mostly of background events. These plots confirm that the model is a close match to the data for each discriminating variable.

Projection plots are limited in that each event is assigned an equal weight, irrespective of the likelihood ratio associated with it. This problem is addressed by the use of sPlots, described in the following Section.

$\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$

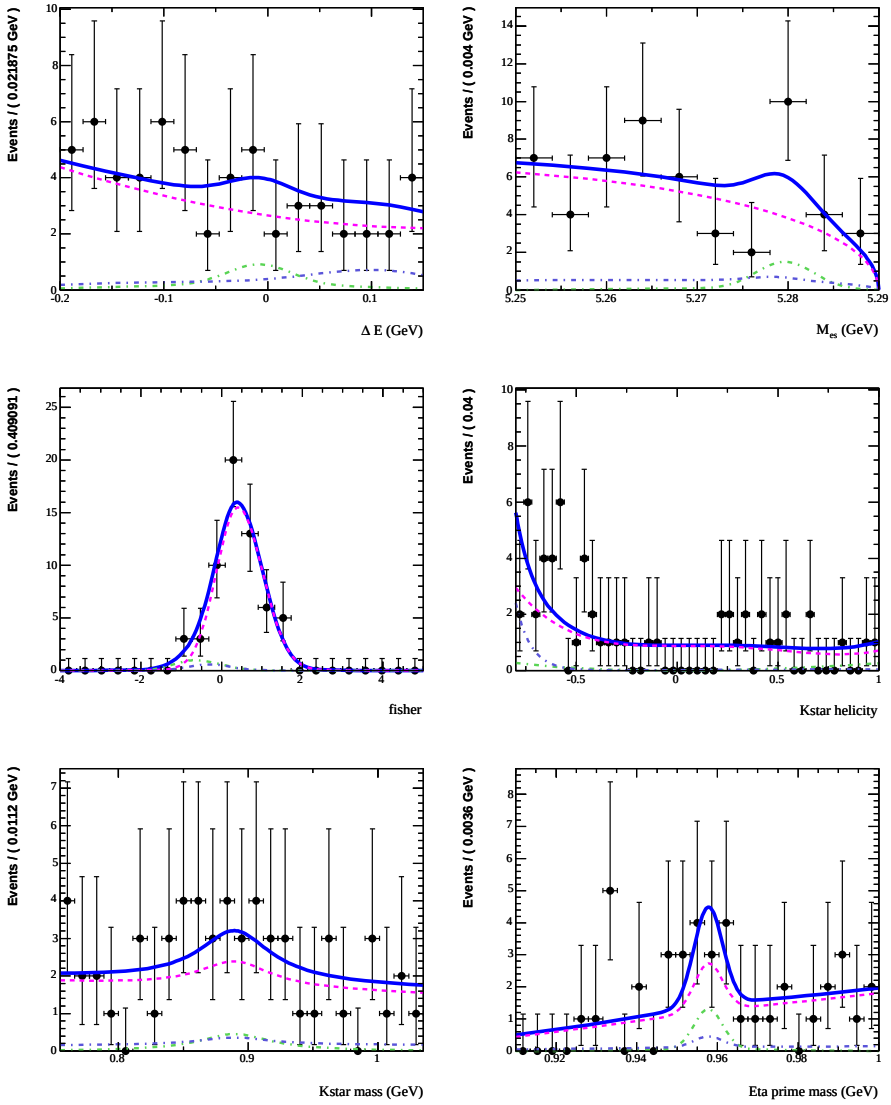

Figure 6.4: Projection plots for $\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$. For each discriminating variable we show the pdf model (solid blue) overlaying the data and the fit components: signal (green), $q\overline{q}$ (magenta) and $B\overline{B}$ (dotted blue).

$\eta'_{\rho\gamma} K^{*+}_{K^+\pi^0}$

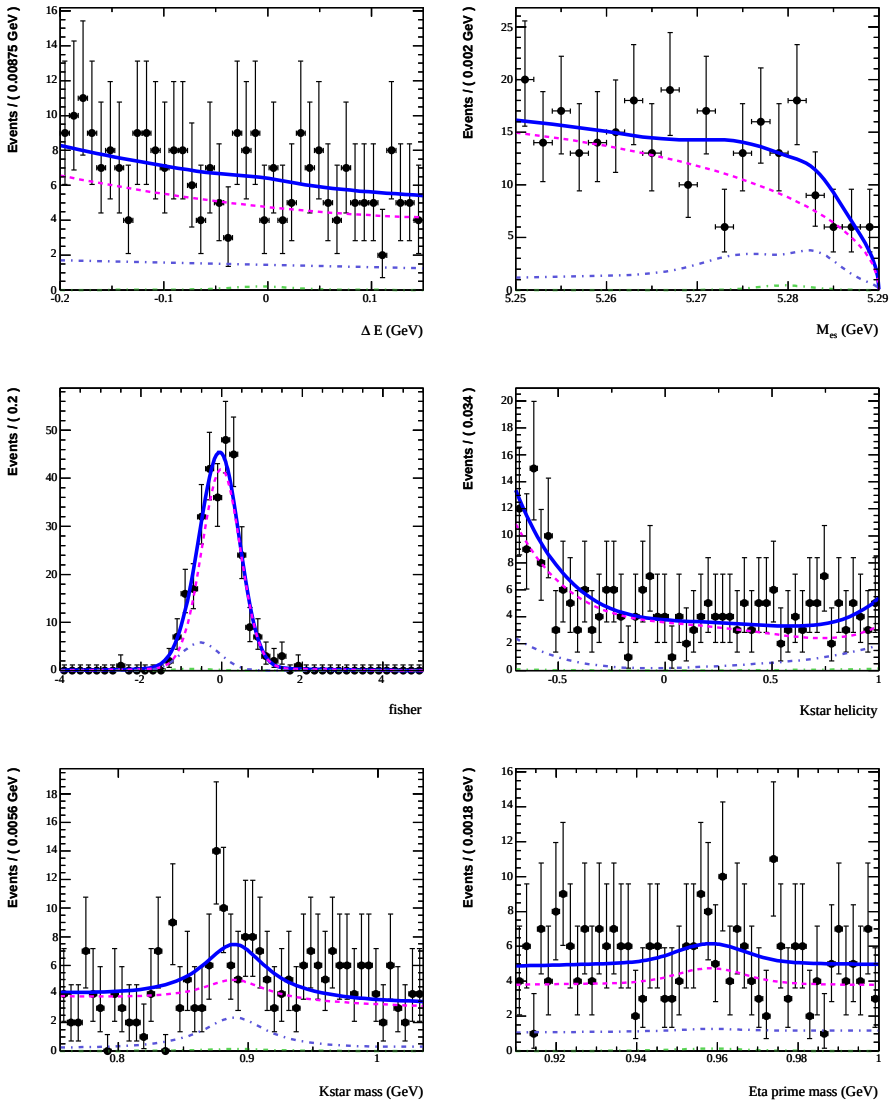

Figure 6.5: Projection plots for $\eta'_{\rho\gamma} K^{*+}_{K^+\pi^0}$. For each discriminating variable we show the pdf model (solid blue) overlaying the data and the fit components: signal (green), $q\overline{q}$ (magenta) and $B\overline{B}$ (dotted blue).

6.4.3 sPlots

Projection plots are limited by the numbers of events which are discarded when performing the cut on the likelihood ratio. Subsequently, it may be difficult to be certain that the

resulting distribution is free from anomalous contributions from either signal or background, or indeed whether any features seen are purely statistical fluctuations. The sPlot technique addresses this shortcoming by keeping all events in the data sample, binning and weighting them according to the PDF of interest and the covariance matrix of the PDF of interest. The sPlot-distribution weights are normalised to the yields in each PDF, and the uncertainties in each bin are simply given by the square root of the weights. Figures 6.6 and 6.7 on the following pages show sPlots [38] for the two principal modes, and Tables 6.4 and 6.5 show the fit yields obtained while making sPlots. sPlots and fit yields for the other four modes used in the combined fit can be found in Section B.5 on page 147.

$\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$

sPlot for	**nBkg (fit)**	**nChls (fit)**	**nSig (fit)**
ΔE	2959 ± 58	53 ± 21	8 ± 6
m_{ES}	2966 ± 57	50 ± 20	3 ± 6
$\mathcal{F}$	2993 ± 62	21 ± 30	7 ± 8
$\cos\theta_H$	2984 ± 60	31 ± 26	6 ± 5
m^{K^*}	2966 ± 57	50 ± 19	3 ± 4
$m^{\eta'}$	2969 ± 57	45 ± 19	5 ± 5

Table 6.4: Fit yields for $q\overline{q}$ background, charmless $B\overline{B}$ background and signal observed while making sPlots for the mode $\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$, effectively removing one discriminating variable at a time from the fit.

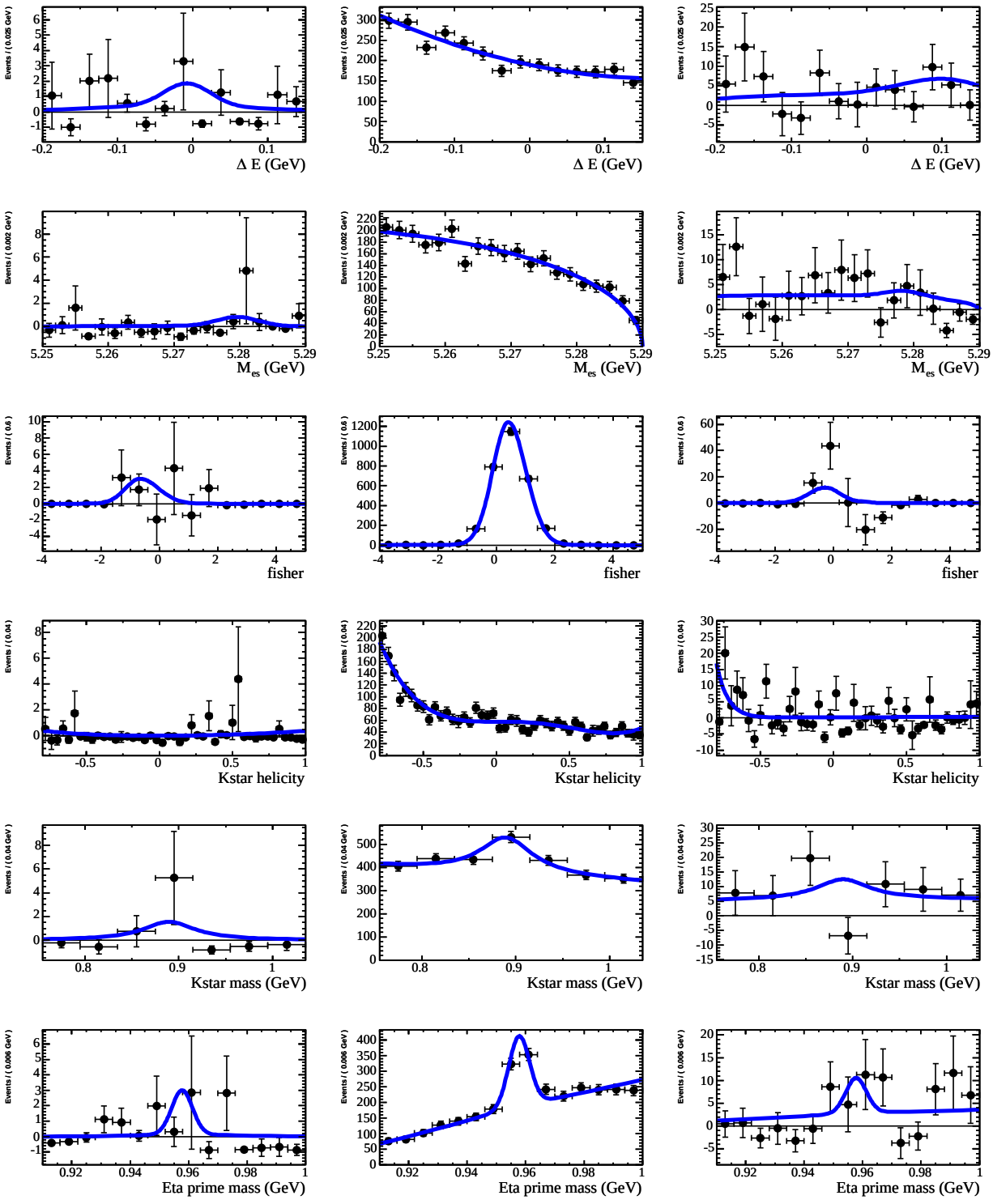

Figure 6.6: sPlots for $\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$. The columns are signal (left), $q\overline{q}$ background (middle), and $B\overline{B}$ background (right).

$\boldsymbol{\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}}$

sPlot for	nBkg (fit)	nChls (fit)	nSig (fit)
ΔE	12668 ± 125	309 ± 63	19 ± 17
m_{ES}	12632 ± 127	367 ± 68	-2 ± 17
$\mathcal{F}$	12631 ± 139	345 ± 89	20 ± 19
$\cos\theta_H$	12655 ± 129	350 ± 69	-8 ± 12
m^{K^*}	12700 ± 128	289 ± 66	7 ± 13
$m^{\eta'}$	12649 ± 125	350 ± 63	-1 ± 16

Table 6.5: Fit yields for $q\overline{q}$ background, charmless $B\overline{B}$ background and signal observed while making sPlots for the mode $\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$, effectively removing one discriminating variable at a time from the fit.

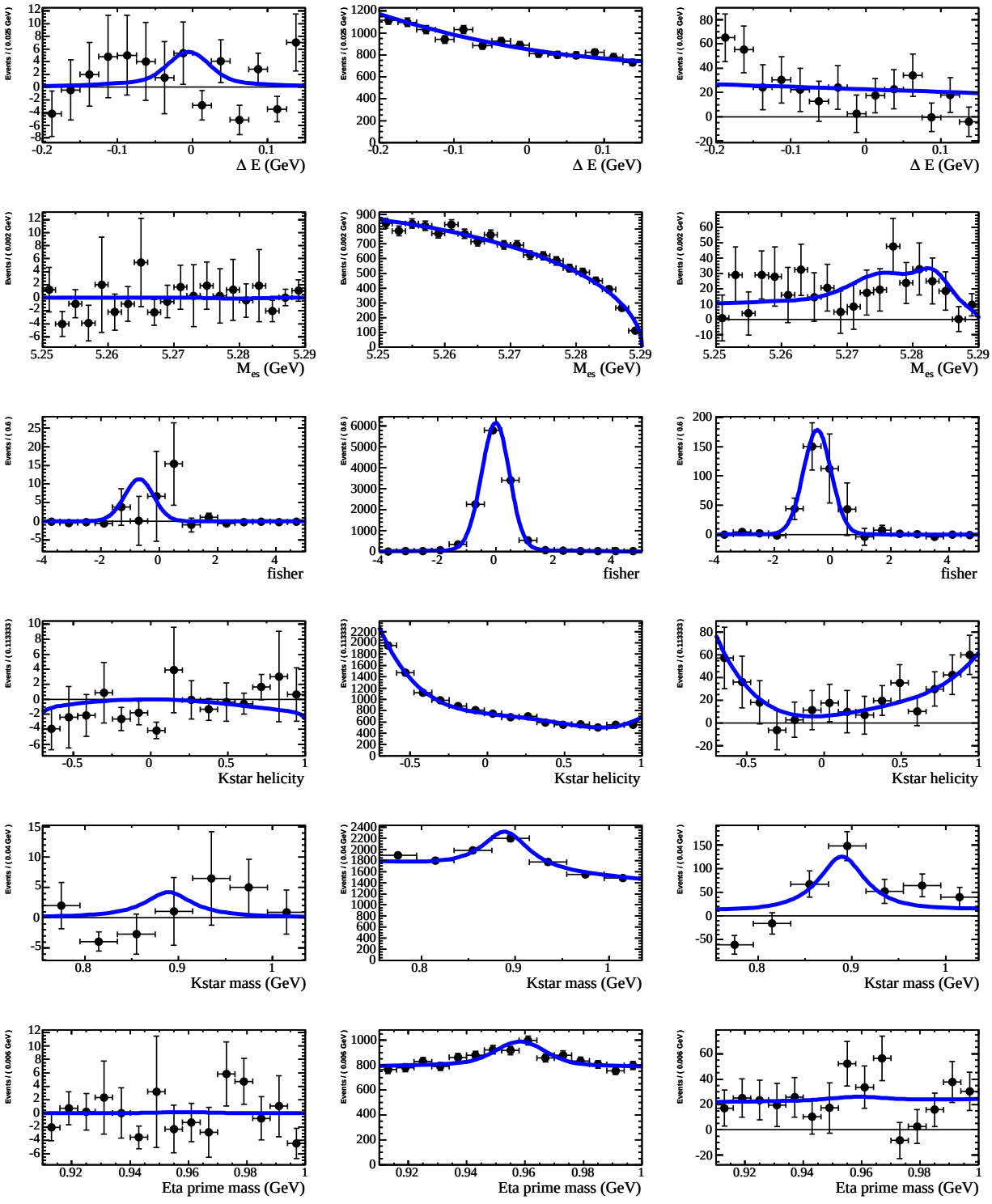

Figure 6.7: sPlots for $\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$. The columns are signal (left), $q\overline{q}$ background (middle), and $B\overline{B}$ background (right).

Chapter 7

Systematic uncertainties

In Table 7.1 the various sources of systematic errors are calculated, and described below. Systematics which are correlated (C) and uncorrelated (U) between different sub-decay modes are distinguished for the branching fractions of interest. The treatment of correlated and uncorrelated systematic errors is different when combining subdecay modes and it is therefore important to differentiate between the two types. It is also important to distinguish between errors which are multiplicative in nature compared to those which should be summed (in quadrature), and these are also shown in the Table.

Track multiplicity

An error of 1% is assigned to reflect small differences between MC and data due to the imposition of the condition that requires at least one charged track from the other B in the event to be reconstructed.

Tracking efficiency

The tracking efficiencies in MC and data are not identical. We apply a 1.3% uncertainty for the η' and K^{*+} daughter pions, and a 1.4% for the daughter pions from the ρ and K^{*0}.

In addition, the decay modes containing a K_S^0 also require an additional factor to reflect the systematic uncertainty arising from the K_S^0 efficiency correction (see Section 3.5.2), which is a further $1.8 - 1.9\%$ applied to each subdecay mode in which a K_S^0 is present.

Neutrals correction

A linear addition of 3% per π^0 or η in the final state is required, to reflect uncertainty in the neutral efficiency. In addition, a further 1.8% per γ from decays $\eta' \to \rho^0\gamma$ is required.

Luminosity, B counting

An error of 1.1% on B counting included from the error given by the official luminosity B counting script is required.

$\cos\theta_{\mathrm{T}}$

An uncertainty related to the $\cos\theta_{\mathrm{T}}$ cut is required since a discrepancy is observed between data and MC, which is proportional to the tightness of the cut. This is discussed in detail in Appendix A, and an uncertainty of between 0.5% and 3% is applied.

Branching fractions of daughters

This is simply the uncertainty on the branching fractions from the PDG [13].

MC statistics

The error in the Monte Carlo events selected is given by the following expression:

$$\frac{\sqrt{\epsilon(1-\epsilon)/N}}{\epsilon} \tag{7.1}$$

where ϵ is the "MC ϵ (%)" listed in Table 6.1 on page 103, and N is the total number of signal MC events generated for each mode, either 124,000 or 125,000, as described in Section 3.2 on page 57.

Signal Model

The parameters describing the signal PDFs are fixed in the nominal fit. The shift and scale factors applied to correct observed differences between data and Monte Carlo are derived from studies of control samples as shown in Table 4.2 on page 73. These are applied to ΔE, m_{ES} and the η' mass in the modes with the decay $\eta' \to \rho^0\gamma$.

In the case of the Fisher discriminant PDF, the signal parameters are varied by uncertainties derived from control samples as discussed in Appendix A and shown in Figure A.3.

Systematic errors for all other signal PDF parameters are determined by changing the PDF parameters by the statistical uncertainties obtained when fitting the PDF shapes to signal MC. Fits are run with the varied parameters, taking correlations into consideration, yielding the systematic errors from the variation in the yields.

ML fit bias

We assign a systematic uncertainty of one-half the fit bias (see Section 5.2), adding the statistical uncertainty from the toys used to measure the bias in quadrature.

Charmless $B\overline{B}$

We float the yields of charmless $B\overline{B}$ background events in each mode; the statistical uncertainty on the signal yield therefore may contain a contribution from the $B\overline{B}$ background. For all modes (except $\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$), the fitted charmless yield is larger than expected. To establish a systematic on this we fix the charmless yield to $n_{B\bar{B},\mathrm{expected}} + \sigma_{B\bar{B},\mathrm{fit}}$ and take the difference between the signal yield obtained in this fit and the nominal fit as an additional, asymmetric systematic uncertainty.

S-wave $B\overline{B}$

We assign a systematic uncertainty due to possible presence of an unaccounted-for non-resonance $K\pi$ S-wave background. We use half the value of the shift in our signal yield as determined by our studies of this background (see Sec. 4.5.3) when the S-wave component is fixed to its expected value.

Charge Asymmetry

We assign a systematic error of 0.02 absolute on our measurements of $\mathcal{A}_{\rfloor\langle}$. The primary sources of systematic uncertainty are due to charged kaon identification (K^{*0} and $K^{*+}_{K^0\pi^+}$ final states) and slow pion reconstruction efficiency ($K^{*+}_{K^0\pi^+}$ final state). This uncertainty is insignificant compared to the statistical uncertainty in these measurements, detailed in Appendix A.

Conclusion

The two principal modes in this analysis show different sources of dominant errors as shown in Table 7.1. The subdecay $\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$ has large systematic uncertainties which are dominated by the large additive contributions from $B\bar{B}$ backgrounds, while the channel $\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$ by contrast is much cleaner, to the extent that small additive fit bias and S-wave systematic error assume prominence. The systematics for all six in the combined fit are dominated by the additive, uncorrelated contributions.

Table 7.1: Systematic uncertainties for all decay modes. For each contribution we show if it is correlated [C] or uncorrelated [U] between the different submodes. For the modes $\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$ and $\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$, an uncertainty of $1.8-1.9\%$ due to the K^0_S efficiency correction has been added linearly to the tracking efficiency (*). The total additive errors are included in the uncorrelated errors on the line below them.

Quantity	$\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$	$\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$	$\eta'_{\eta\pi\pi}K^{*+}_{K^+\pi^0}$	$\eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$	$\eta'_{\eta\pi\pi}K^{*0}$	$\eta'_{\rho\gamma}K^{*0}$
Multiplicative errors (%)						
Track multiplicity [C]	1.0	1.0	1.0	1.0	1.0	1.0
Tracking efficiency [C]	5.9 *	6.1 *	3.9	4.1	5.4	5.6
$\pi^0/\eta_{\gamma\gamma}/\gamma$ eff [C]	3.0	1.8	6.0	4.8	3.0	1.8
Number $B\overline{B}$ [C]	1.1	1.1	1.1	1.1	1.1	1.1
$\cos\theta_T$ [C]	0.5	1.3	0.5	1.3	0.5	3.0
Branching fractions [U]	3.4	3.4	3.4	3.4	3.4	3.4
MC statistics [U]	0.6	0.6	0.8	0.9	0.6	0.6
Total multiplicative (%)	7.6	7.5	8.1	7.5	7.3	7.6
Total multiplicative [$\pm\mathcal{B}(10^{-6})$]	0.47	0.35	0.23	0.22	0.30	0.25
Additive errors (events)						
Signal model [U]	0.35	1.1	0.13	1.6	0.55	0.83
Fit bias [U]	0.45	1.5	0.54	1.3	0.90	4.8
$B\overline{B}$ background [U]	0.2	1.9	0.3	17	0.82	5.2
S-wave background [U]	0.7	1.2	0.8	2.0	1.8	4.8
Total additive (events)	$^{+0.92}_{-0.90}$	$^{+2.9}_{-2.2}$	$^{+1.0}_{-0.97}$	$^{+17}_{-2.6}$	$^{+2.2}_{-2.1}$	$^{+8.6}_{-6.8}$
Total errors [$\mathcal{B}(10^{-6})$]						
Total Additive	$^{+0.55}_{-0.54}$	$^{+1.2}_{-0.88}$	$^{+0.71}_{-0.68}$	$^{+9.5}_{-1.4}$	$^{+0.44}_{-0.41}$	$^{+1.13}_{-0.89}$
Uncorrelated	$^{+0.59}_{-0.58}$	$^{+1.2}_{-0.90}$	$^{+0.72}_{-0.68}$	$^{+9.5}_{-1.4}$	$^{+0.46}_{-0.43}$	$^{+1.13}_{-0.90}$
Correlated	±0.42	±0.31	±0.27	±0.19	±0.26	±0.22

Chapter 8

Conclusions

8.1 Results in light of theoretical predictions

Theoretical predictions from two different approaches, and the (as of 2005) experimental status was presented in Table 1.2 on page 31, and are reproduced below. Also included is the the neutral mode, the two subdecays of which are used in the combined six-mode $B \to \eta' K^*$ fit. The central value of the updated combined charged mode agrees very well with the central values predicted by QCD factorisation while still being within errors of the SU(3) flavour predictions. For the mode $B^0 \to \eta' K^{*0}$ a branching fraction which is consistent with both theoretical approaches is observed. The QCD factorisation numbers are subject to uncertainties which mean it is difficult to draw any firm conclusions regarding the effectiveness of the approach.

The results presented here were published in 2007 [39]. Shortly after publication, the Belle collaboration published an updated search for the same final states, but failed to find signal in either the charged or neutral channels [40]. The analysis was based on a sample of 535 million $B\bar{B}$ pairs, well in excess of *BABAR*'s data-set. The published 90% confidence upper limits are below the central values measured in the analysis presented here. The results are in poor agreement, if not inconsistent.

The Belle analysis imposed much tighter cuts on many of the analysis variables, e.g., a $2.5\,\sigma$ mass window for η', corresponding to $15\,\mathrm{MeV/c^2}$, compared to the $90\,\mathrm{MeV/c^2}$ range of

Mode	Theory		Experimental results			**Updated results [39]**		
	SU(3) [18]	QCD [19]	HFAG	*BABAR* [16]	Belle [20]	Value	σ	U.L.
$\eta' K^{*0}$	$3.0^{+1.2}_{-0.3}$	$3.9^{+9.2}_{-5.1}$	< 7.6	< 7.6	< 20	$3.8 \pm 1.1 \pm 0.5$	$4.3\,\sigma$	–
$\eta' K^{*+}$	$2.8^{+1.2}_{-0.3}$	$5.1^{+10.3}_{-\ 5.9}$	< 14	< 14	< 90	$4.9^{+1.9}_{-1.7} \pm 0.8$	$3.6\,\sigma$	< 7.9

Table 8.1: Comparison of the combined results for both charged and neutral channels to theoretical predictions and experimental status at time of publication in 2007. Note branching fractions are all quoted in units of ($\times 10^{-6}$), and all upper limits are quoted at 90% confidence. The final column lists updated upper limits (U.L.) as appropriate.

acceptance used here. Belle imposed a minimum photon energy of 100 MeV for photons originating from π^0 decays, whereas this analysis imposes a cut below 30 MeV. The consequence of imposing these tight cuts appears to be borne by the efficiency numbers, with Belle's combined corrected efficiencies a factor of 2 or 3 lower than this analysis. It is therefore probable that Belle's cuts disposed of a significant number of signal events that would be included if our cuts were adopted. Also, Belle's systematic errors are often much larger than those presented in this analysis. The K_s reconstruction systematic is 4 %, compared to 1.9% in this analysis. The result is that there are large uncertainties in Belle's measured signal yields.

8.1.1 Subsequent developments

The *BABAR* collaboration published an update to this analysis in 2010 [41], using an expanded data-set comprising some 426 fb^{-1} of integrated luminosity, corresponding to 467 million $B\bar{B}$ pairs. The results of this, along with the Belle analysis, and updated theoretical predictions are summarised in Table 8.2. The updated *BABAR* results are in excellent agreement with those published here, and also with an updated SU(3) flavour symmetry prediction [42]. Clearly these are also in poor agreement with the results from Belle.

The ratio of branching fractions for $B^+ \to \eta' K^{*+}$ and $B^+ \to \eta K^{*+}$ seen with the data presented in this analysis and elsewhere, is consistent with predictions outlined in Section

Mode	Theoretical predictions		Experimental results				
	SU(3) [42]	QCD f. [19]	*BABAR* [39]	Belle [40]	**Run1-6 *BABAR* [41]**		
$B^0 \to \eta' K^{*0}$	4.2 ± 1.6	$3.9^{+9.2}_{-5.1}$	$3.8 \pm 1.1 \pm 0.5$	< 2.6	$3.1^{+0.9}_{-0.8} \pm 0.3$	4σ	–
$B^+ \to \eta' K^{*+}$	4.3 ± 1.6	$5.1^{+10.3}_{-\ 5.9}$	$4.9^{+1.9}_{-1.7} \pm 0.8$	< 2.9	$4.8^{+1.6}_{-1.4} \pm 0.8$	3.8σ	< 7.2

Table 8.2: Comparison of the combined results for both charged and neutral modes for analyses completed after the one presented here. The subsequent *BABAR* analysis is in good agreement with the analysis presented in this thesis; both of these are in poor agreement with the updated 90% confidence limits from Belle, whose analysis did not identify signal in either decay channel. Note that branching fractions are all quoted in units of ($\times 10^{-6}$), and all upper limits are quoted at 90% confidence.

1.4.1 which give credence to either a non-standard treatment of mixing in the $\eta - \eta'$ system or the inclusion of other contributions e.g. from OZI-suppressed electroweak penguins. Lipkin's calculation [15] of the effect of such a treatment on the ratio of branching fractions for $B^+ \to \eta K^{*+}$ and $B^+ \to \eta' K^{*+}$ is supported by the latest HFAG data, since the current branching fraction for $B^+ \to \eta K^{*+}$ is 19.3 ± 1.6 (both expressed in units of $\times 10^{-6}$). The ratio of the central values of η to η' decays involving a K^* is $4.0^{+1.6}_{-1.4}$, consistent with Lipkin's estimation of 5.6 (no errors are quoted). QCD factorisation [19] results in a ratio of 2.1 ± 0.53 while a flavour SU(3) treatment [18] produces in a ratio of $11.4^{+10.9}_{-3.6}$. All of these values are in stark contrast to the ratio of 33 estimated by Lipkin (no errors are quoted) using a standard treatment of $\eta' - \eta$ mixing, and hence a more sophisticated treatment is warranted.

Appendix A

$B^- \to D^0\pi^-$ control sample studies

A.1 Data vs Monte Carlo comparison for $B^- \to D^0\pi^-$

A.1.1 Motivation

It is imperative that the Monte Carlo events generated for use in this analysis are a close match to the data obtained from the experiment. The parameters of the signal MC PDFs are reliant on the generated events, and any discrepancies will introduce unwanted uncertainties when these parameters are used in the final fits to the data. By choosing to analyse a sample from a decay which is similar to the studied modes, any such discrepancies can be quantified and corrected in the signal MC PDFs. The D^0 decay channel studied is $D^0 \to K\pi\pi^0$, which when combined with the charged pion closely is almost identical in composition to the final states analysed here. It is reasonable to expect that the bahaviour of MC versus on-peak data will closely mirror the modes under study.

This control sample is used in several ways. We measure and correct differences in m_{ES}, ΔE event variables. We also use the sample to quantify systematics on the measurement of charge asymmetry, study the efficiency of our $\cos\theta_{\mathrm{T}}$ cuts, and finally use the sample to examine the accuracy of the Fisher parameterisation.

A.1.2 Data-sets

For the MC sample we have taken 4.4M events produced according to a cocktail of $B^+ \to D^{(*)0}\ \pi$ decays. Although the data-set is only skimmed for $B^+ \to D^0$ final states, often the slow pion from D^{*0} decays is lost, resulting in backgrounds from D^{*0}. We fit for this later to account for reflections in the ΔE distribution.

A.1.3 Event selection

The same event selection criteria is applied to the data and MC samples outlined above. These cuts are applied in a similar fashion to Section 3.

- $|\cos\theta_{\rm T}| \leq 0.75$,
- $|\Delta E| \leq 0.3\,{\rm GeV}$,
- $N_{\rm trks} \geq 4$,
- $5.20 \leq {\rm m}_{ES} \leq 5.29$,
- $-3 < \mathcal{F} < 4$, where $\mathcal{F}$ is the Legendre-based Fisher,
- $1845 < m^{D^0}_{K^-\pi^+\pi^0} < 1885\ \ {\rm MeV}/c^2$ (about a 2 sigma cut),
- $120 < m^{\pi^0}_{\gamma\gamma} < 150\ \ {\rm MeV}/c^2$,
- For the π^0 daughters, $E_\gamma > 30$ MeV.

A.1.4 Results for $\mathbf{m}_{ES}$ and $\boldsymbol{\Delta E}$

Final results for ${\rm m}_{ES}$ and ΔE are obtained by requiring $|\Delta E| < 0.06$ GeV and $5.274 < {\rm m}_{ES} < 5.286$ respectively.

The ${\rm m}_{ES}$ distributions are fitted with a Gaussian for the signal and an ARGUS function for the background. For ΔE we fit using two Gaussians, one centred around zero for the

signal D^0 decays and another at approximately -200 MeV to represent background from D^{*0} decays. The remaining background is fitted using a 2nd order polynomial. Single Gaussians are used throughout for fitting the signal to best facilitate any necessary scaling between data and MC.

In Table A.1 the results are summarised for the mean and sigma of the signal ΔE and m_{ES} distributions, and plots showing ΔE and m_{ES} plots for both MC and data are displayed in Figures A.2 and A.1. The indicated corrections to the MC are listed at the bottom of the Section of Table A.1, and these are in good agreement with the actual corrections applied, detailed in Section 4.3 on page 72.

Table A.1: Measured parameters for ΔE and m_{ES} in $D^0\pi$ control sample. All units in MeV.

Sample	$< \Delta E >$	$\sigma_{\Delta E}$	$< \mathrm{m}_{ES} >$	$\sigma_{m_{ES}}$
Combined Sample				
Data	-4.0 ± 0.9	24.8 ± 0.3	5279.60 ± 0.03	2.88 ± 0.03
MC	-2.8 ± 0.1	23.5 ± 0.1	5279.21 ± 0.01	2.77 ± 0.01
Adjust. to MC	$-(1.2 \pm 1.0)$	$\times(1.06 \pm 0.02)$	$+(0.4 \pm 0.01)$	$\times(1.04 \pm 0.01)$

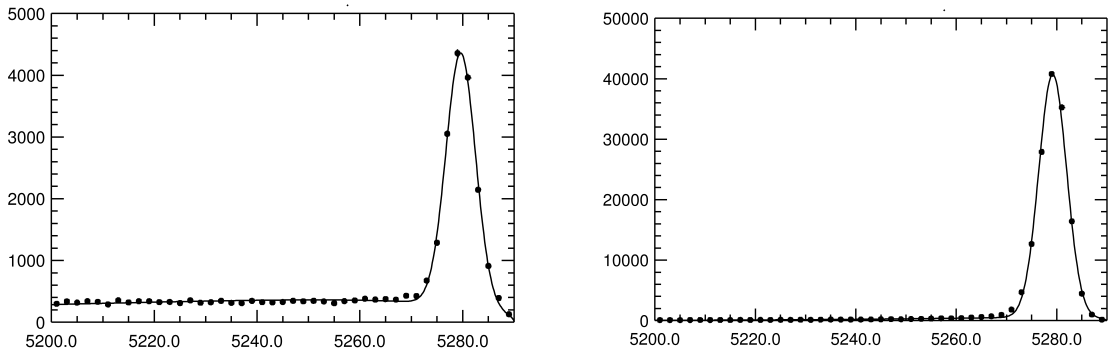

Figure A.1: m_{es} distributions for data (left) and Monte Carlo (right) with overlaid fit for $B^- \to D^0\pi^-$ control sample studies.

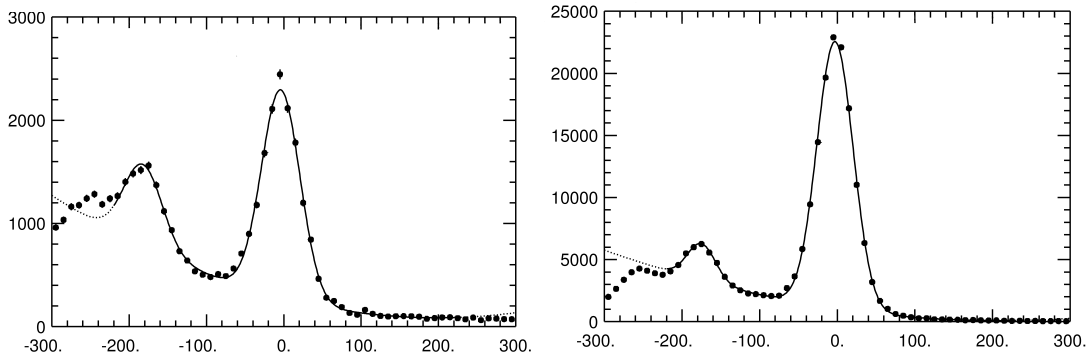

Figure A.2: ΔE distributions for data (left) and Monte Carlo (right) with overlaid fit for $B^- \to D^0\pi^-$ control sample studies. The ΔE plots for data and MC show reflections from D^{*0} mesons being misreconstructed as D^0 mesons where the slow π has been lost from the D^{*0}. The low side peak is around 170 MeV/c^2, corresponding to the pion mass (130 MeV/c^2) plus the difference in mass between the misreconstructed D^{*0} and the D^0 (circa 40 MeV/c^2).

A.1.5 Charge asymmetry systematic errors

If we take the m_{ES} distribution and split it into positive and negative B charges then this allows us to study the charge asymmetry in this mode. Fitting the signal with a single Gaussian as before and the background with an ARGUS function then we can extract the number of signal events for B^+ and B^- and calculate the asymmetry

$$\frac{n_{B^-} - n_{B^+}}{n_{B^-} + n_{B^+}} \tag{A.1}$$

A.1.6 Charge asymmetry results

The asymmetry histogram of the m_{ES} distributions for the full control sample are shown in Figure A.3. The full data sample contains 7286 ± 97 B^- events and 7459 ± 99 B^+ events. The charge asymmetry for the signal is $\mathcal{A}_{D\pi}^{Data} = -0.0118 \pm 0.0133$, and is dominated by the known asymmetry associated with the charged kaon selector. The asymmetry in the MC sample is $\mathcal{A}_{D\pi}^{MC} = -0.0102 \pm 0.0040$. The asymmetry in signal events is in good agreement with the systematic of 1.1% which we have associated with this in Section 7.

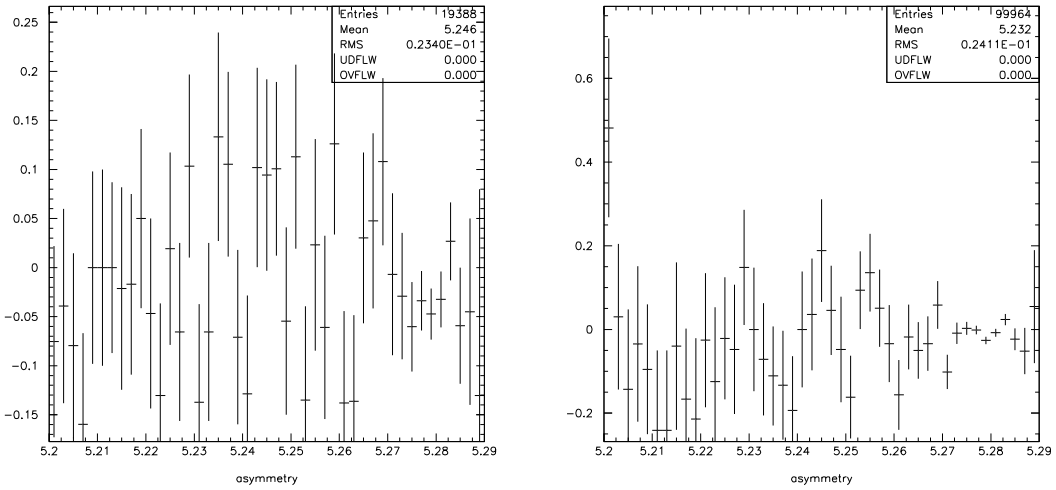

Figure A.3: m_{ES} charge asymmetry distribution for data (left) and MC (right).

A.1.7 Measurement of $\cos\theta_T$ efficiency

To study whether or not the MC accurately models the thrust variable, we apply successively tighter cuts to the data and MC samples and compare the rate at which events are lost. The ratio of this rate in data and MC should be one. The cut on the cosine of the thrust angle is tightened from 0.9 to 0.5 in steps of 0.1. All other selection criterion remain unchanged from above. The results in Table A.2 are from the full data-set only. These clearly indicate a small deviation between data and MC, which gets worse as the thrust cut is tightened. This motivates the requirement for a systematic which is dependent on the cut applied. The systematic uncertainties presented in Table 7.1 are confirmed as reasonable by examining the final column in Table A.2.

A.1.8 Study of Fisher discriminant and its parameterisation

We use the large signal statistics in the control sample to study the distribution of the Fisher variable and its parameterisation in data and MC. We first fit the signal MC Fisher distribution with an asymmetric Gaussian and determine its parameters. The asymmetric Gaussian parameterisation consists of the mean and RMS of the distribution, and the asymmetry $\mathcal{A}$ between the right and left Gaussian widths (σ). Second, we run the full ML fit on the real

Table A.2: Results of $\cos\theta_{\rm T}$ cut study. The final column demonstrates that the MC's modelling of signal is accurate to within a few percent, dependent on the magnitude of the $\cos\theta_{\rm T}$ applied. Similar corrections are therefore justified in the main analysis.

$\cos\theta_{\rm T}$ cut	data events	MC events	MC/data	MC/data normalised
0.9	18550 ± 166	171907 ± 437	9.28 ± 0.11	1
0.8	15963 ± 147	150134 ± 407	9.41 ± 0.11	1.014 ± 0.024
0.7	13588 ± 132	129124 ± 377	9.50 ± 0.12	1.024 ± 0.025
0.6	11347 ± 119	108923 ± 347	9.60 ± 0.13	1.034 ± 0.026
0.5	9228 ± 106	89189 ± 314	9.67 ± 0.15	1.042 ± 0.029

data, while allowing the signal Fisher parameters to float in the fit. The results obtained in MC and real data for the mean, RMS, and the asymmetry $\mathcal{A}$ are listed in Table A.3. The sys-

Table A.3: Fisher PDF parameters for the $B^- \to D^0\pi^-$ control sample.

Variable	signal MC	real data	Δ(data-MC)	Systematic
Mean μ	-0.6280 ± 0.0014	-0.6271 ± 0.0044	0.0009 ± 0.0046	± 0.005
RMS	0.5126 ± 0.0010	0.4987 ± 0.0032	-0.0139 ± 0.0034	± 0.014
Asymmetry $\mathcal{A}$	0.1237 ± 0.0037	0.0611 ± 0.0140	-0.0626 ± 0.0145	± 0.064

tematic uncertainty on each parameter is determined as the quadratic sum of the data-MC difference and its own statistical uncertainty. We find the systematics are $\sigma_{\rm syst}(\mu) = \pm 0.005$, $\sigma_{\rm syst}(\text{RMS}) = \pm 0.014$, and $\sigma_{\rm syst}(A) = \pm 0.064$.

Appendix B

PDFs, projection plots, sPlots and fit validation tables

This Appendix contains many plots relevant to the four modes used in the combined overall fit result which are not the two principal modes presented in the main body of the thesis. These are found from the pages below:

B.1 PDF plots

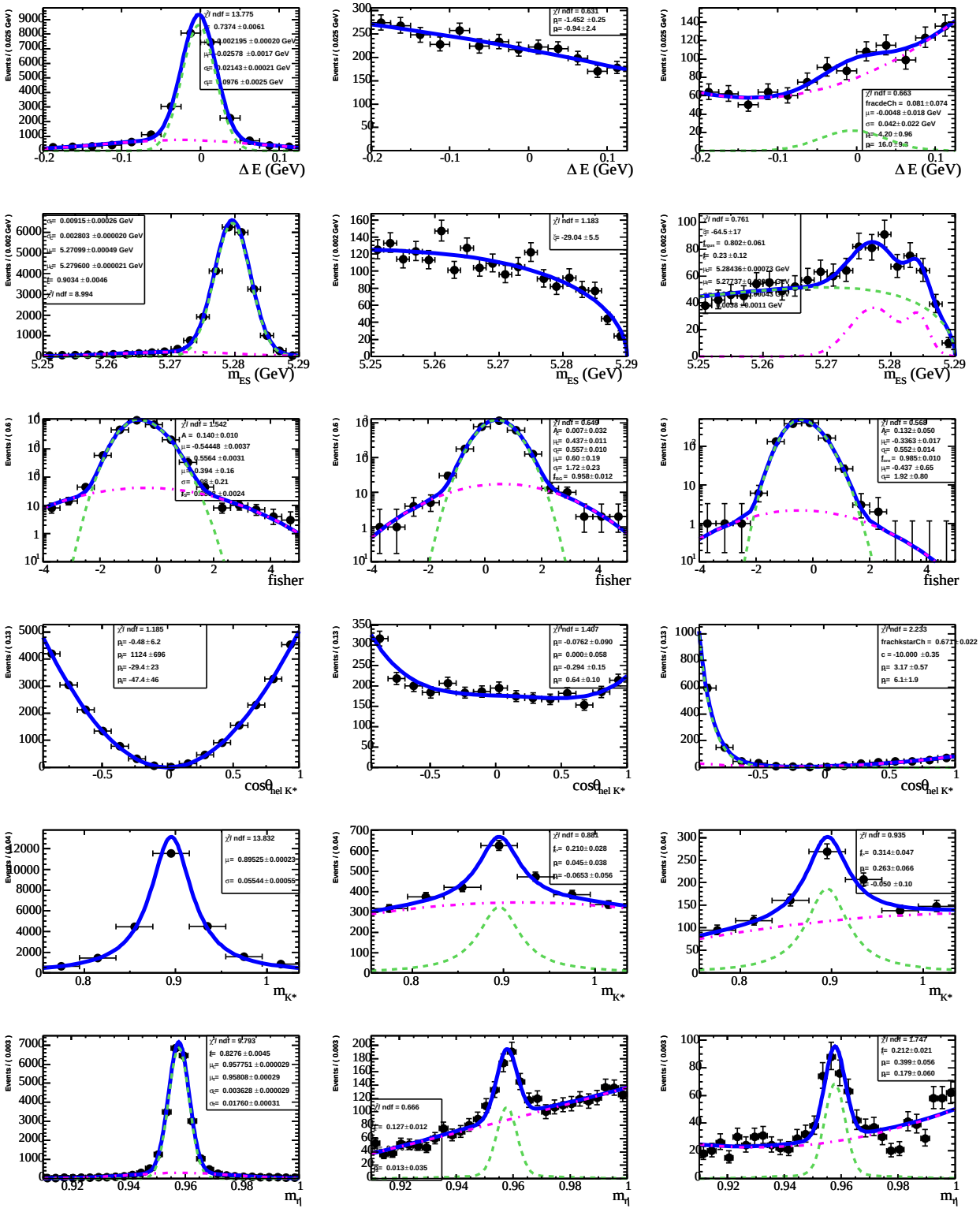

Figure B.1: PDFs for the mode $\eta'_{\eta\pi\pi}K^{*0}$: Signal MC (*left*), on-peak sidebands (*centre*) and $B\overline{B}$ background (*right*). From top to bottom: ΔE, m_{ES}, $\mathcal{F}$, $\cos\theta_{\mathcal{H}}$, $m_{K^\star}$, $m_{\eta'}$. The solid blue curve is the sum of the dotted components described in Table 4.1.

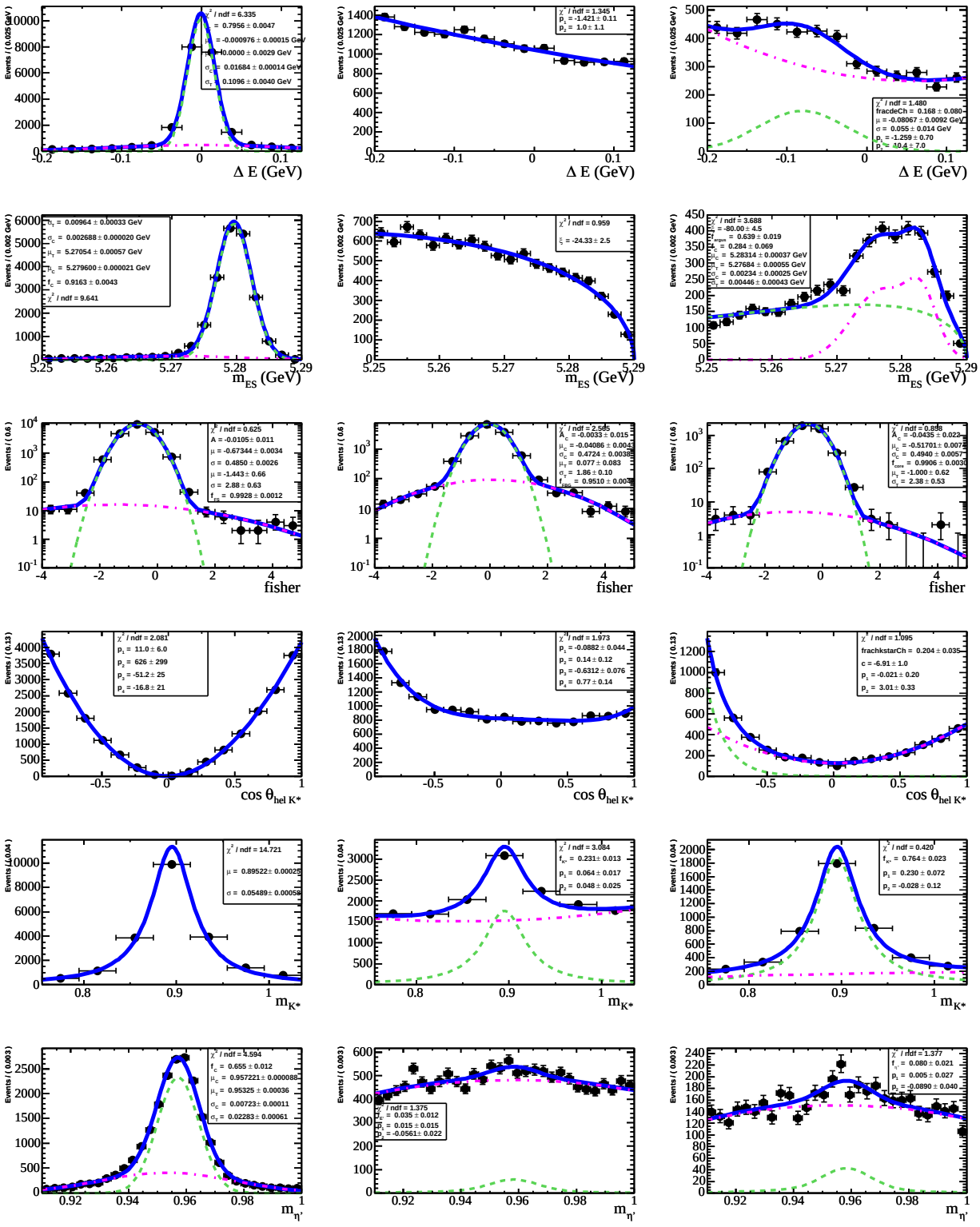

Figure B.2: PDFs for the mode $\eta'_{\rho\gamma}K^{*0}$: Signal MC (*left*), on-peak sidebands (*centre*) and $B\overline{B}$ background (*right*). From top to bottom: ΔE, m_{ES}, $\mathcal{F}$, $\cos\theta_{\mathcal{H}}$, $m_{K^\star}$, $m_{\eta'}$. The solid blue curve is the sum of the dotted components described in Table 4.1.

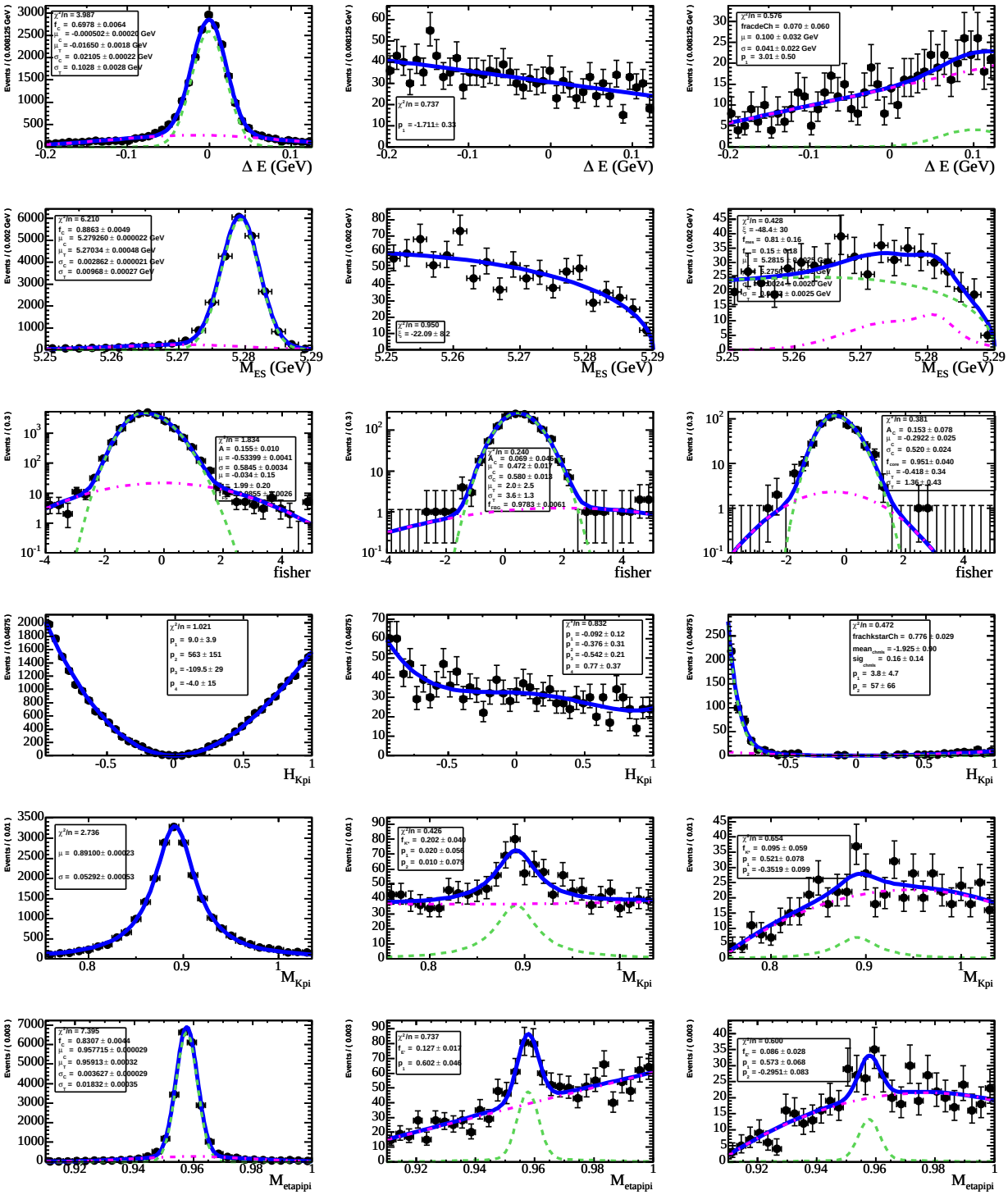

Figure B.3: PDFs for the mode $\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$: Signal MC (*left*), on-peak sidebands (*centre*) and B background (*right*). From top to bottom: ΔE, m_{ES}, $\mathcal{F}$, $\cos\theta_{\mathcal{H}}$, $m_{K^\star}$, $m_{\eta'}$. The solid blue curve is the sum of the dotted components described in Table 4.1.

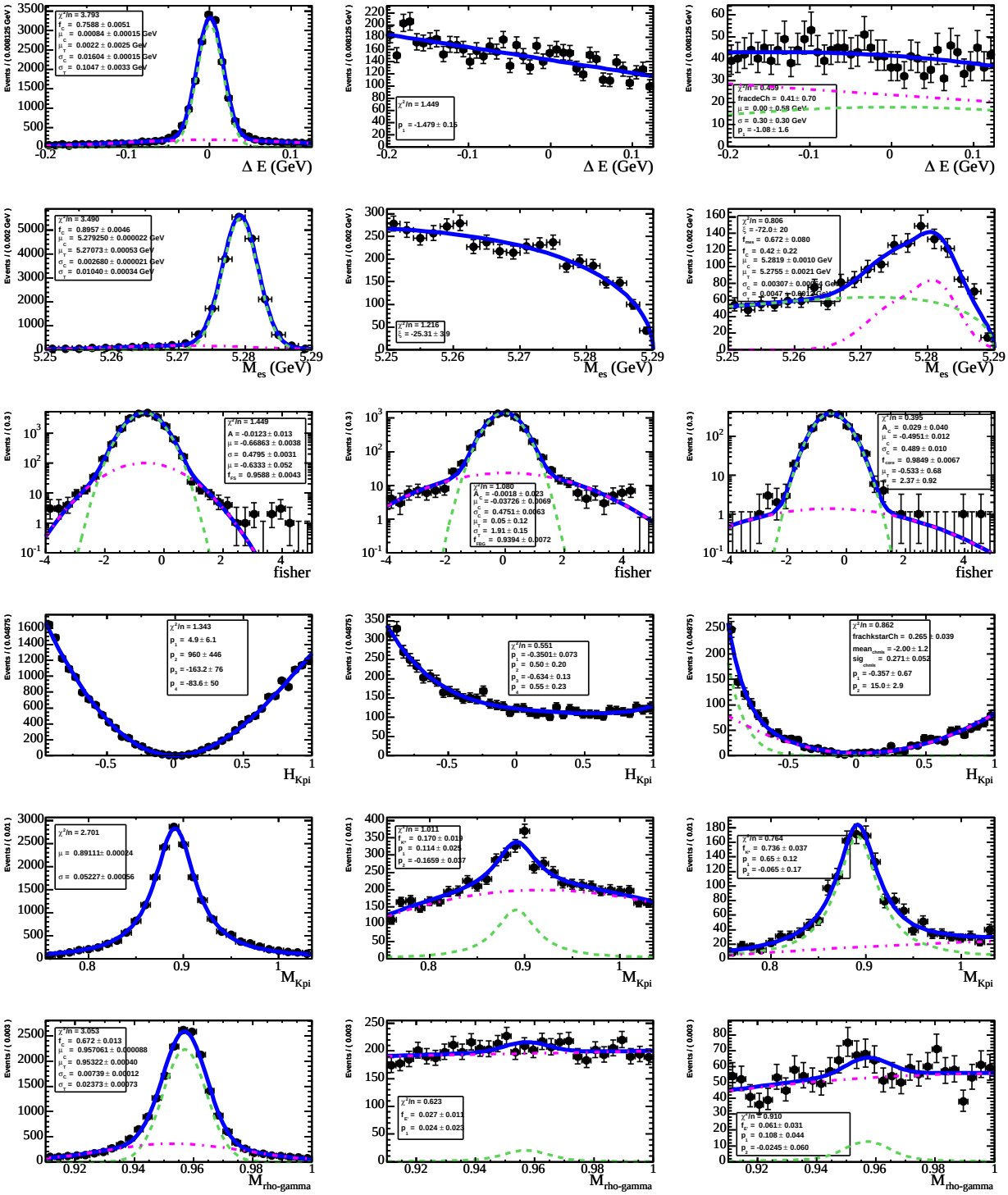

Figure B.4: PDFs for the mode $\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$: Signal MC (*left*), on-peak sidebands (*centre*) and B background (*right*). From top to bottom: ΔE, m_{ES}, $\mathcal{F}$, $\cos\theta_{\mathcal{H}}$, $m_{K^\star}$, $m_{\eta'}$. The solid blue curve is the sum of the dotted components described in Table 4.1.

B.2 Charmless $B\bar{B}$ details

Table B.1: Charmless $B\overline{B}$ backgrounds to the $\eta'_{\eta\pi\pi}K^{*0}_{K\pi}$ mode. For each background channel considered, we list the decay channel including all subdecays, production mode number, the selection efficiency (MC ϵ), an estimate for the branching fraction ($\mathcal{B}$), the daughter product branching fraction $\Pi\,\mathcal{B}_i$ from the PDG and used in the MC production, the expected contribution of this channel in the on-peak data sample (Norm. #) and the number of simulated events used in the charmless cocktail (# in PDF).

Signal mode Bkg. channel	Mode #	MC ϵ (%)	Est. $\mathcal{B}$ (10^{-6})	$\prod\mathcal{B}_i$	Norm. # $B\overline{B}$ Bkg.	# in PDF Bkg. file
$\eta'_{\eta\pi\pi}K^{*0}_{K\pi}$						
$B^+ \to \eta'_{\eta_{\gamma\gamma}\pi\pi}K^+$	1506	0.46	69	0.174	12.8	658
$B^0 \to \eta_{3\pi}K^{*0}_{K^+\pi^-}$	1540	0.7	19	0.151	4.7	240
$B^0 \to \eta'_{\eta_{\gamma\gamma}\pi\pi}K_S$	1510	0.32	63	0.060	2.8	146
$B^0 \to\ a_1^0K^{*0}(L, f_L = 0.7)$	5329	0.03	21	0.467	0.7	38
$B^0 \to \omega\ K^{*0}_{K^+\pi^-}(L, f_L = 1)$	2507	0.11	4	0.594	0.6	31
$B^+ \to\ a_1^+(\rho^+\pi^0)K^{*0}(L, f_L = 0.7)$	5331	0.02	42	0.233	0.4	21
Total					22	1134

Table B.2: Charmless $B\overline{B}$ backgrounds to the $\eta'_{\rho\gamma}K^{*0}_{K\pi}$ mode. For each background channel considered, we list the decay channel including all subdecays, production mode number, the selection efficiency (MC ϵ), an estimate for the branching fraction ($\mathcal{B}$), the daughter product branching fraction $\Pi\,\mathcal{B}_i$ from the PDG and used in the MC production, the expected contribution of this channel in the on-peak data sample (Norm. #) and the number of simulated events used in the charmless cocktail (# in PDF).

Signal mode Bkg. channel	Mode #	MC ϵ (%)	Est. $\mathcal{B}$ (10^{-6})	$\prod\mathcal{B}_i$	Norm. # $B\overline{B}$ Bkg.	# in PDF Bkg. file
$\eta'_{\rho\gamma}K^{*0}_{K\pi}$						
$B^0 \to\ a_1^0 K^{*0}(L, f_L = 0.7)$	5329	1.4	21	0.467	31.8	776
$B^+ \to\ a_1^+(\rho^0\pi^+)K^{*0}(L, f_L = 0.7)$	5333	1.3	42	0.233	29.5	720
$B^0 \to \phi_{3\pi}K^{*0}_{K^+\pi^-}(T, f_L = 0.50)$	4975	6.62	9.5	0.103	15	366
$B^0 \to \phi_{3\pi}K^{*0}_{K^+\pi^-}(L, f_L = 0.50)$	4974	5.93	9.5	0.103	13.5	328
$B^+ \to\ a_1^+(\rho^+\pi^0)K^{*0}(L, f_L = 0.7)$	5331	0.51	42	0.233	11.7	284
$B^0 \to\ a_1^0 K^{*0}(T, f_L = 0.7)$	5330	1.19	21	0.200	11.5	282
$B^+ \to \eta'_{\rho\gamma}K^+$	1508	0.23	69	0.295	10.6	259
$B^+ \to\ a_1^+(\rho^0\pi^+)K^{*0}(T, f_L = 0.7)$	5334	1.08	42	0.100	10.5	255
$B^+ \to \rho^+\ K^{*0}_{K^+\pi^-}(L, f_L = 0.7)$	2244	0.78	7	0.667	8.5	207
$B^+ \to \rho^+\ K^{*0}_{K^+\pi^-}(T, f_L = 0.7)$	2243	1.08	3	0.667	5	122
$B^0 \to \eta'_{\rho\gamma}K_S$	1511	0.24	63	0.101	3.6	88
$B^+ \to\ a_1^0 K^+$	4874	0.24	6	1.000	3.3	81
$B^0 \to \rho^0 K^{*0}_{K^+\pi^-}(L, f_L = 0.5)$	2359	0.86	2.5	0.667	3.3	80
$B^+ \to\ a_1^+(\rho^0\pi^+)\rho^0(L, f_L = 1)$	4105	0.06	48	0.500	3.3	80
$B^0 \to\ a_1^-(\rho^0\pi^-)K^+$	4871	0.21	13	0.500	3.1	76
$B^0 \to \rho^0 K^{*0}_{K^+\pi^-}(T, f_L = 0.5)$	2360	0.75	2.5	0.667	2.9	71
$B^+ \to\ a_1^+(\rho^+\pi^0)K^{*0}(T, f_L = 0.7)$	5332	0.25	42	0.100	2.5	60
$B^+ \to\ a_1^0 K^{*+}(K^+\pi^0)(L, f_L = 1)$	5327	0.14	21	0.333	2.3	55
$B^0 \to\ a_1^-(\rho^0\pi^-)K^{*+}(K^+\pi^0)((L, f_L = 1)$	5325	0.13	42	0.167	2.1	51
$B^+ \to \phi_{3\pi}K^+$	2713	0.65	9	0.155	2.1	51
$B^0 \to \omega\ K^{*0}_{K^+\pi^-}(L, f_L = 1)$	2507	0.37	4	0.594	2	50
$B^0 \to\ a_1^-(\rho^-\pi^0)K^+$	4960	0.13	13	0.500	1.9	47
$B^+ \to \eta'_{\rho\gamma}K^{*+}_{K^+\pi^0}$	2773	1.4	6	0.098	1.9	46
$B^0 \to \eta_{3\pi}K^{*0}_{K^+\pi^-}$	1540	0.25	19	0.151	1.7	41
$B^0 \to\ a_1^+(\rho^0\pi^+)\rho^-(L, f_L = 1)$	4002	0.02	84	0.500	1.6	38
$B^+ \to\ a_1^0\rho^+(L, f_L = 1)$	3999	0.01	48	1.000	1.5	36
$B^+ \to\ a_1^+(\rho^+\pi^0)\rho^0(L, f_L = 1)$	4107	0.02	48	0.500	1.1	27
$B^0 \to\ a_1^-(\rho^-\pi^0)K^{*+}(K^+\pi^0)(L, f_L = 1)$	5323	0.06	42	0.167	1	25
$B^0 \to \phi_{K^+K^-}\eta'_{\rho\gamma}$	2929	3.02	1	0.145	1	24
$B^+ \to \omega\ K^+$	1250	0.08	5	0.891	0.8	19
$B^0 \to \rho^- K^{*+}_{K^+\pi^0}(T, f_L = 0.25)$	2500	0.06	9	0.333	0.4	10
Total					191	4655

Table B.3: Charmless $B\overline{B}$ backgrounds to the mode $\eta'_{\eta\pi\pi}K^{*+}_{K_S\pi^+}$. For each background channel considered, we list the decay channel including all subdecays, production mode number, the selection efficiency (MC ϵ), an estimate for the branching fraction ($\mathcal{B}$), the daughter product branching fraction $\Pi\,\mathcal{B}_i$, the expected contribution of this channel in the on-peak data sample (Norm. #) and the number of simulated events used in the charmless cocktail (# in PDF).

Signal mode Bkg. channel	Mode #	MC ϵ (%)	Est. $\mathcal{B}$ (10^{-6})	$\prod\mathcal{B}_i$	Norm. # $B\overline{B}$ B kg.	# in PDF Bkg. file
$\eta'_{\eta\pi\pi}K^{*+}_{K_S\pi^+}$						
$B^0 \to \eta'_{\eta_{\gamma\gamma}\pi\pi}K_S$	1510	0.92	63	0.060	8.1	410
$B^+ \to \eta_{3\pi}K^{*+}_{K_S\pi^+}$	1537	0.67	24	0.052	1.9	98
$B^+ \to\ a_1^0K^{*+}_{K_S\pi^+}(L, f_L = 1)$	6657	0.03	21	0.229	0.3	15
$B^0 \to\ a_1^-(\rho^-\pi^0)K^{*+}_{K_S\pi^+}(L, f_L = 1)$	6659	0.02	42	0.115	0.3	13
$B^0 \to\ a_1^+(\rho^+\pi^0)K_S$	4952	0.01	13	0.172	0.1	2
Total					10.7	540

Table B.4: Charmless $B\overline{B}$ backgrounds to the mode $\eta'_{\rho\gamma}K^{*+}_{K_S\pi^+}$. For each background channel considered, we list the decay channel including all subdecays, production mode number, the selection efficiency (MC ϵ), an estimate for the branching fraction ($\mathcal{B}$), the daughter product branching fraction $\Pi\,\mathcal{B}_i$ from the PDG and used in the MC production, the expected contribution of this channel in the on-peak data sample (Norm. #) and the number of simulated events used in the charmless cocktail (# in PDF).

Signal mode Bkg. channel	Mode #	MC ϵ (%)	Est. $\mathcal{B}$ (10^{-6})	$\prod\mathcal{B}_i$	Norm. # $B\overline{B}$ B kg.	# in PDF Bkg. file
$\eta'_{\rho\gamma}K^{*+}_{K_S\pi^+}$						
$B^+ \to\ a_1^0K^{*+}_{K_S\pi^+}(L, f_L = 1)$	6657	1.36	21	0.229	15.1	370
$B^0 \to\ a_1^-(\rho^0\pi^-)K^{*+}_{K_S\pi^+}(L, f_L = 1)$	6661	1.26	42	0.115	14.1	344
$B^0 \to \eta'_{\rho\gamma}K_S$	1511	0.44	63	0.101	6.6	160
$B^0 \to\ a_1^-(\rho^-\pi^0)K^{*+}_{K_S\pi^+}(L, f_L = 1)$	6659	0.52	42	0.115	5.8	142
$B^0 \to \rho^-K^{*+}_{K_S\pi^+}(T, f_L = 0.25)$	2502	1.16	9	0.229	5.5	135
$B^+ \to \phi_{3\pi}K^{*+}_{K_S\pi^+}$	3994	5.97	10	0.035	4.8	118
$B^+ \to \rho^0K^{*+}_{K_S\pi^+}(L, f_L = 1)$	2357	0.79	11	0.229	4.6	113
$B^+ \to\ a_1^+(\rho^0\pi^+)K_S$	4959	0.34	13	0.172	1.7	42
$B^0 \to \rho^-K^{*+}_{K_S\pi^+}(L, f_L = 0.25)$	2501	0.82	3	0.229	1.3	31
$B^0 \to \phi_{3\pi}K_S$	2714	1.3	8	0.053	1.3	31
$B^0 \to\ a_1^0(\rho^+\pi^-)K_S$	4956	0.47	6	0.172	1.1	27
$B^0 \to\ a_1^0(\rho^-\pi^+)K_S$	4955	0.45	6	0.172	1.1	26
$B^0 \to\ a_1^+(\rho^+\pi^0)K_S$	4952	0.21	13	0.172	1.1	26
$B^+ \to \omega\ K^{*+}_{K_S\pi^+}(L, f_L = 1)$	2505	0.41	4	0.204	0.8	18
$B^+ \to\ f_0\ (\pi^+\pi^-)K^{*+}_{K_S\pi^+}$	4623	0.45	5	0.148	0.8	18
$B^0 \to \omega\ K_S$	1536	0.15	6	0.306	0.7	16
$B^0 \to\ a_1^+(\rho^0\pi^+)\rho^-(L, f_L = 1)$	4002	0.01	84	0.500	0.5	12
$B^+ \to\ a_1^0\rho^+(L, f_L = 1)$	3999	0	48	1.000	0.3	7
$B^+ \to\ a_1^+(\rho^0\pi^+)\rho^0(L, f_L = 1)$	4105	0.01	48	0.500	0.3	7
$B^0 \to\ a_1^+(\rho^+\pi^0)\rho^-(L, f_L = 1)$	4001	0	84	0.500	0.1	3
$B^+ \to\ a_1^+(\rho^+\pi^0)\rho^0(L, f_L = 1)$	4107	0	48	0.500	0.1	3
$B^+ \to\ a_1^+(\rho^0\pi^+)K^{*0}(L, f_L = 0.7)$	5333	0	42	0.233	0.1	2
$B^0 \to\ a_1^0K^{*0}(L, f_L = 0.7)$	5329	0	21	0.467	0.1	1
$B^+ \to\ a_1^+(\rho^+\pi^0)K^{*0}(L, f_L = 0.7)$	5331	0	42	0.233	0.1	1
Total					68	1653

B.3 Likelihood ratio distributions

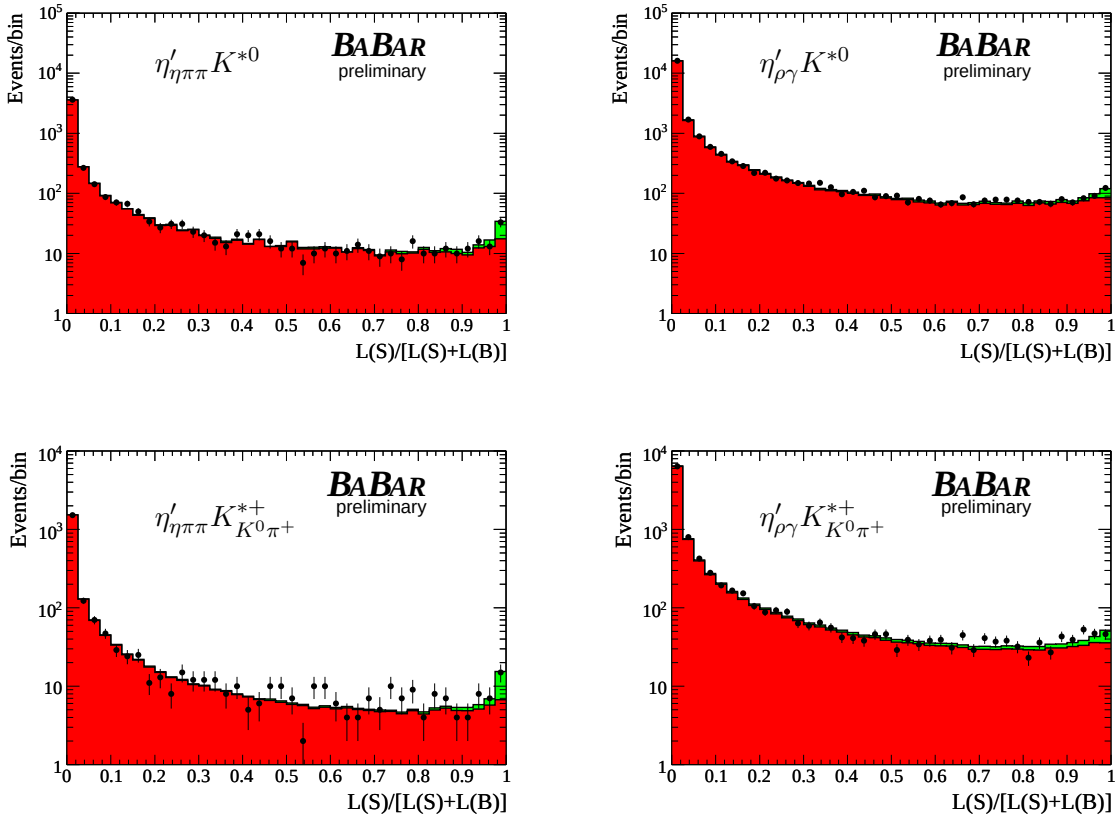

Figure B.5: Plots of the likelihood ratio $\mathcal{L}_{sig}/[\mathcal{L}_{sig} + \sum \mathcal{L}_{bkg}]$ for all modes. The points are the on-resonance data, the red histograms are background toy ($q\overline{q}$ & $B\overline{B}$) and the green histograms are signal toy. The on-resonance points are in good agreement with the toy histograms.

B.4 Projection plots

The following pages display projection plots of the data over each of the fit variables, onto which the result of the combined fit has been overlaid. These plots are obtained using a likelihood ratio $\mathcal{L}_{sig}/[\mathcal{L}_{sig} + \sum \mathcal{L}_{bkg}]$ cut of 0.9 on the data. These plots confirm that the model is a close match to the data for each discriminating variable.

B.4.1 $\eta'_{\eta\pi\pi}K^{*0}$

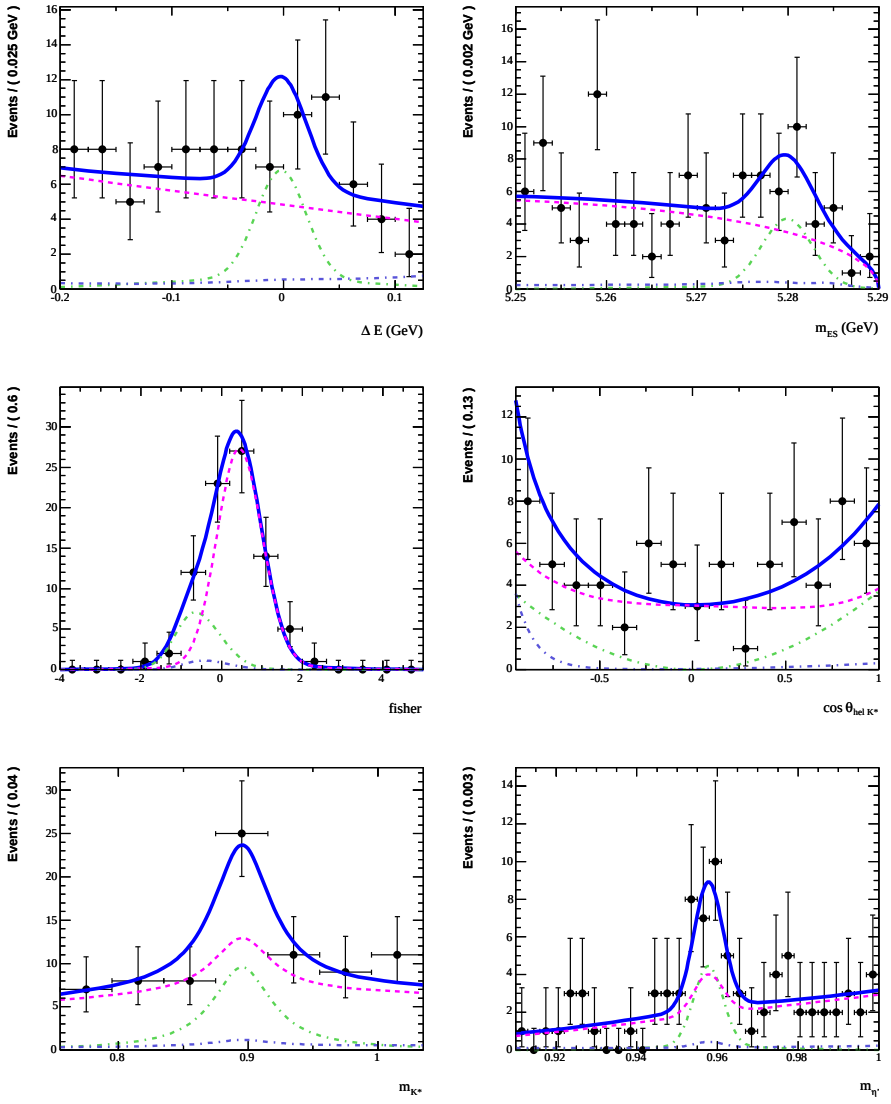

Figure B.6: Projection plots for $\eta'_{\eta\pi\pi}K^{*0}$. For each discriminating variable we show the pdf model (solid blue) overlaying the data and the fit components: signal (green), $q\overline{q}$ (magenta) and $B\overline{B}$ (dotted blue).

B.4.2 $\eta'_{\rho\gamma}K^{*0}$

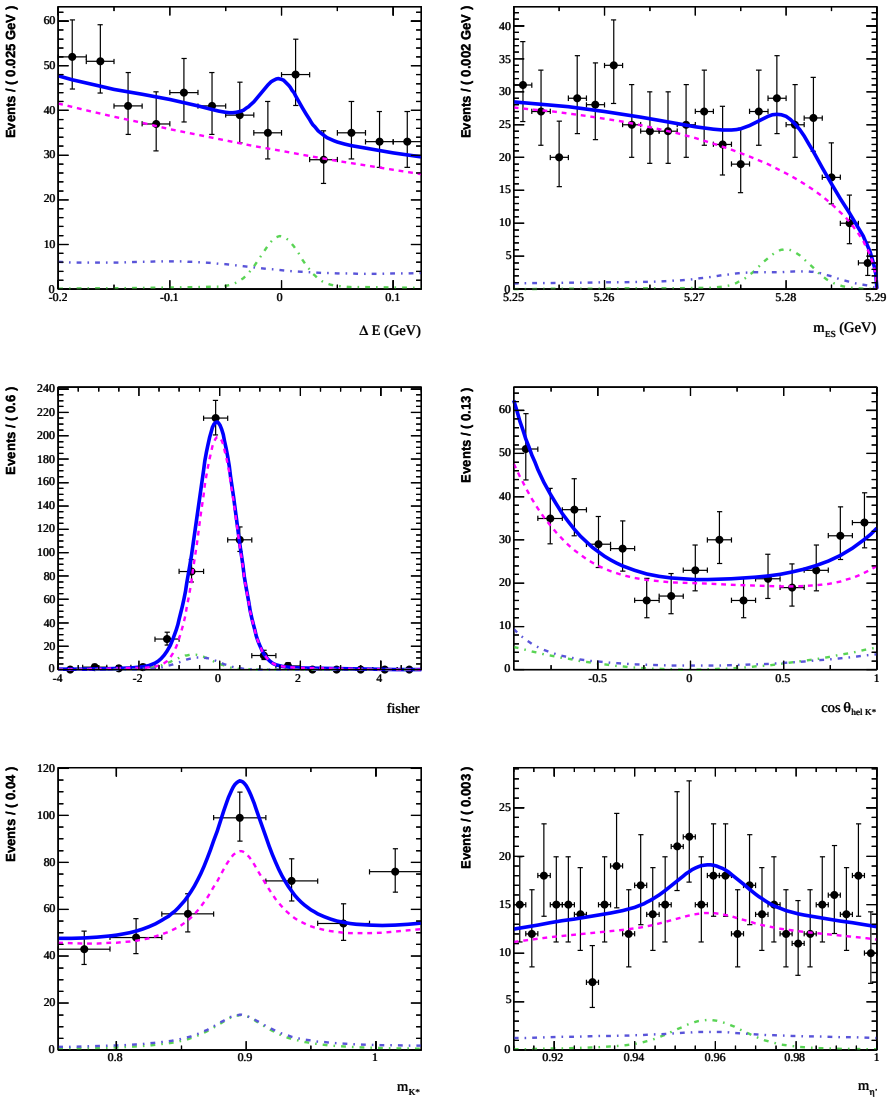

Figure B.7: Projection plots for $\eta'_{\rho\gamma}K^{*0}$. For each discriminating variable we show the pdf model (solid blue) overlaying the data and the fit components: signal (green), $q\overline{q}$ (magenta) and $B\overline{B}$ (dotted blue).

B.4.3 $\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$

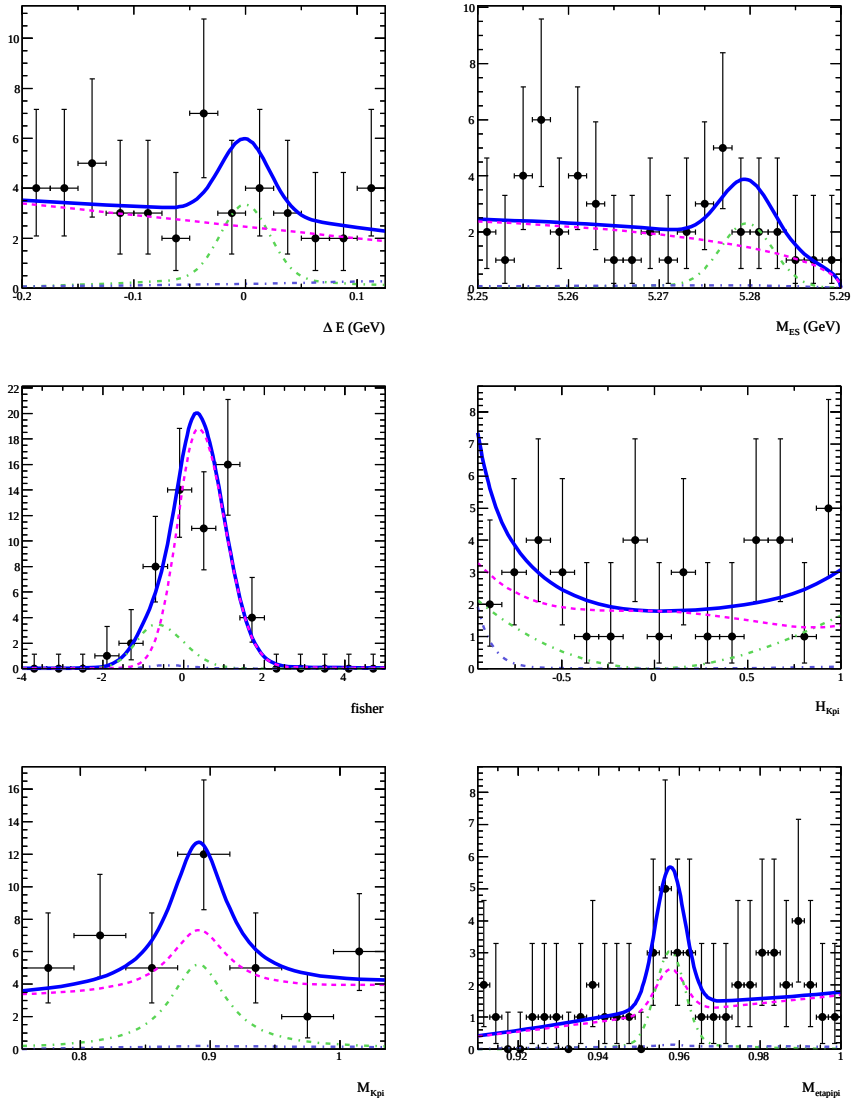

Figure B.8: Projection plots for $\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$. For each discriminating variable we show the pdf model (solid blue) overlaying the data and the fit components: signal (green), $q\overline{q}$ (magenta) and $B\overline{B}$ (dotted blue).

B.4.4 $\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$

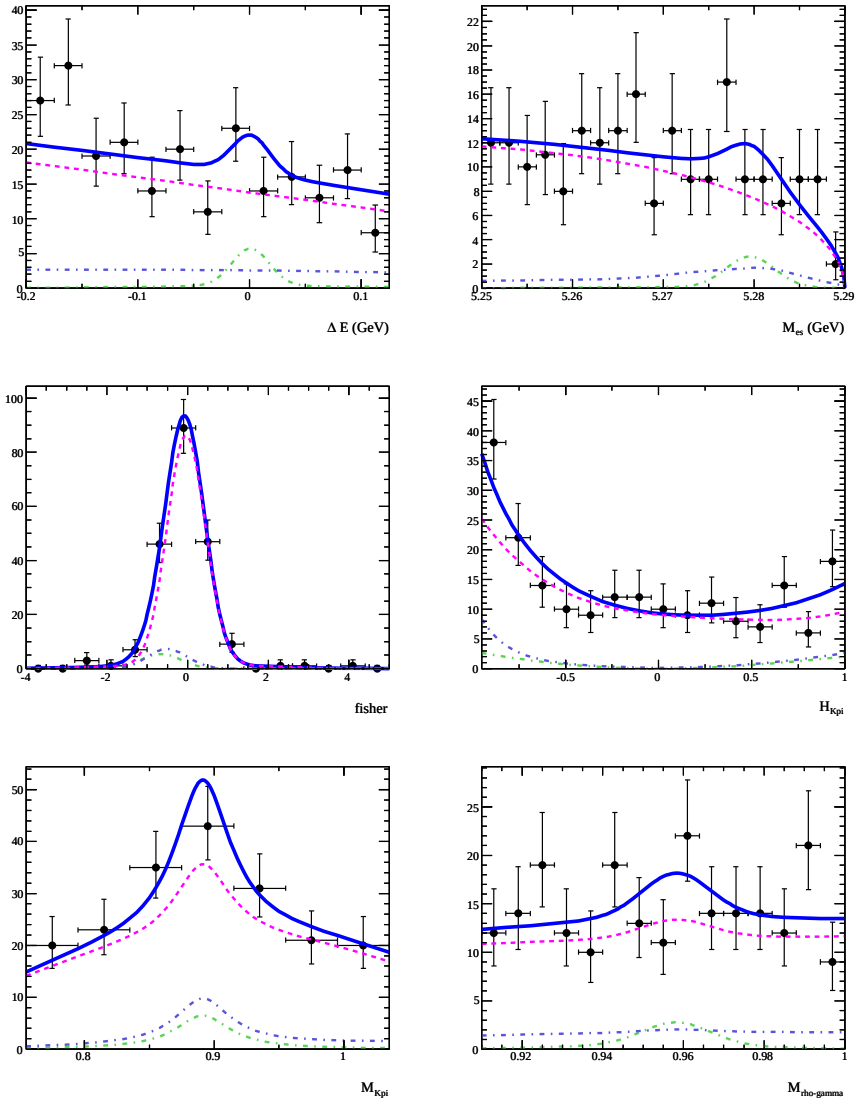

Figure B.9: Projection plots for $\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$. For each discriminating variable we show the pdf model (solid blue) overlaying the data and the fit components: signal (green), $q\overline{q}$ (magenta) and $B\overline{B}$ (dotted blue).

B.5 sPlots

The following pages show sPlots [38] for each mode. The shape of the blue curve in each plot is described in Table 4.1.

B.5.1 $\eta'_{\eta\pi\pi}K^{*0}$

sPlot for	nBkg (fit)	nChls (fit)	nSig (fit)
ΔE	4773 ± 72	37 ± 22	27 ± 9
m_{ES}	4756 ± 72	60 ± 22	21 ± 10
$\mathcal{F}$	4772 ± 78	43 ± 36	22 ± 10
$\cos\theta_H$	4792 ± 75	21 ± 31	24 ± 8
m^{K^*}	4773 ± 72	43 ± 21	21 ± 7
$m^{\eta'}$	4779 ± 72	43 ± 21	15 ± 8

Table B.5: Fit yields for $q\overline{q}$ background, charmless $B\overline{B}$ background and signal observed while making sPlots for the mode $\eta'_{\eta\pi\pi}K^{*0}$, effectively removing one discriminating variable at a time from the fit.

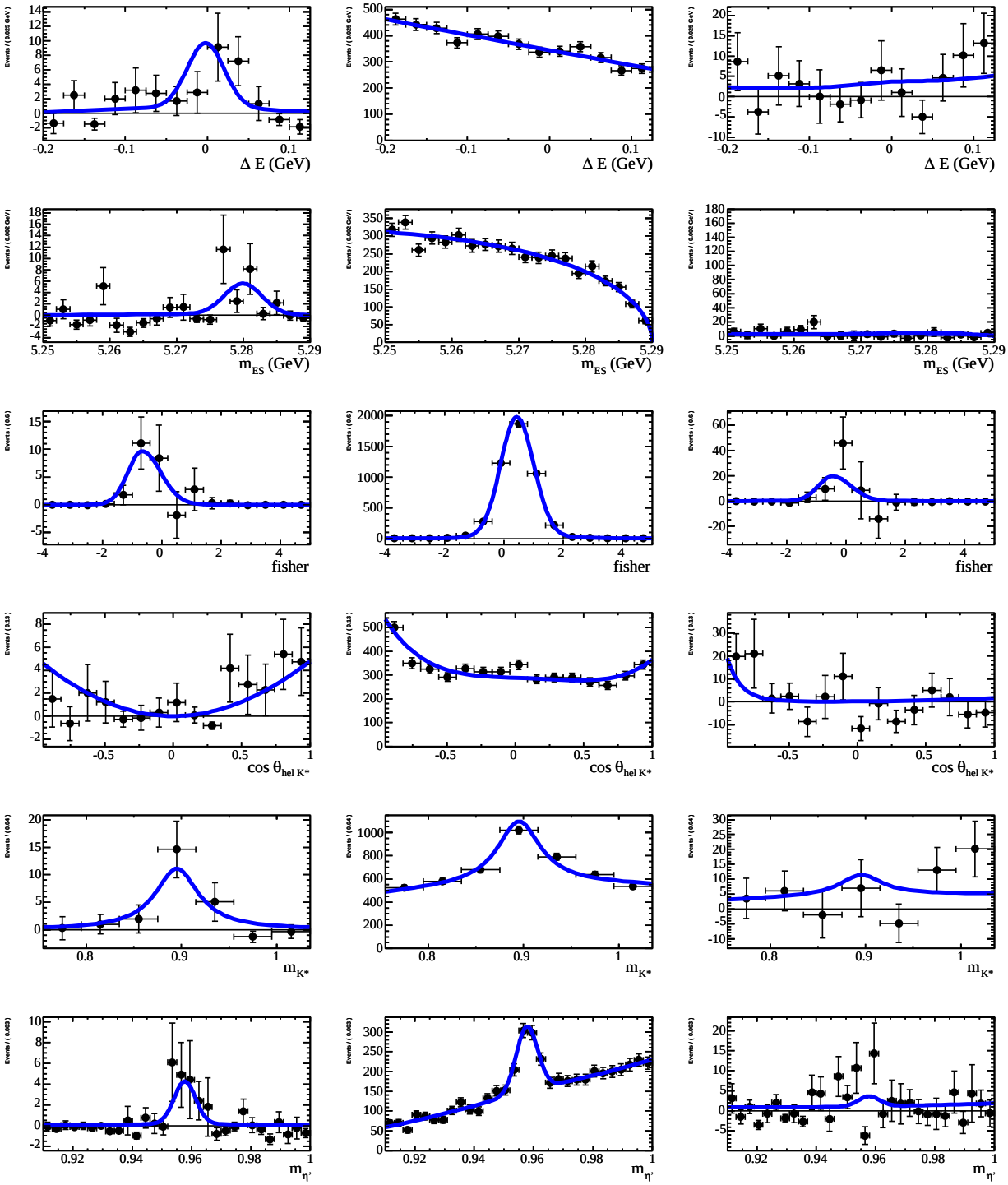

Figure B.10: sPlots for $\eta'_{\eta\pi\pi}K^{*0}$. The columns are signal (left), $q\bar{q}$ background (middle), and $B\overline{B}$ background (right).

B.5.2 $\eta'_{\rho\gamma}K^{*0}$

sPlot for	nBkg (fit)	nChls (fit)	nSig (fit)
ΔE	23384 ± 174	353 ± 95	53 ± 25
m_{ES}	23342 ± 180	426 ± 102	22 ± 20
$\mathcal{F}$	23676 ± 210	70 ± 148	44 ± 19
$\cos\theta_H$	23399 ± 176	357 ± 92	35 ± 15
m^{K^*}	23156 ± 179	600 ± 100	34 ± 15
$m^{\eta'}$	23363 ± 172	388 ± 86	40 ± 17

Table B.6: Fit yields for $q\overline{q}$ background, charmless $B\overline{B}$ background and signal observed while making sPlots for the mode $\eta'_{\rho\gamma}K^{*0}$, effectively removing one discriminating variable at a time from the fit.

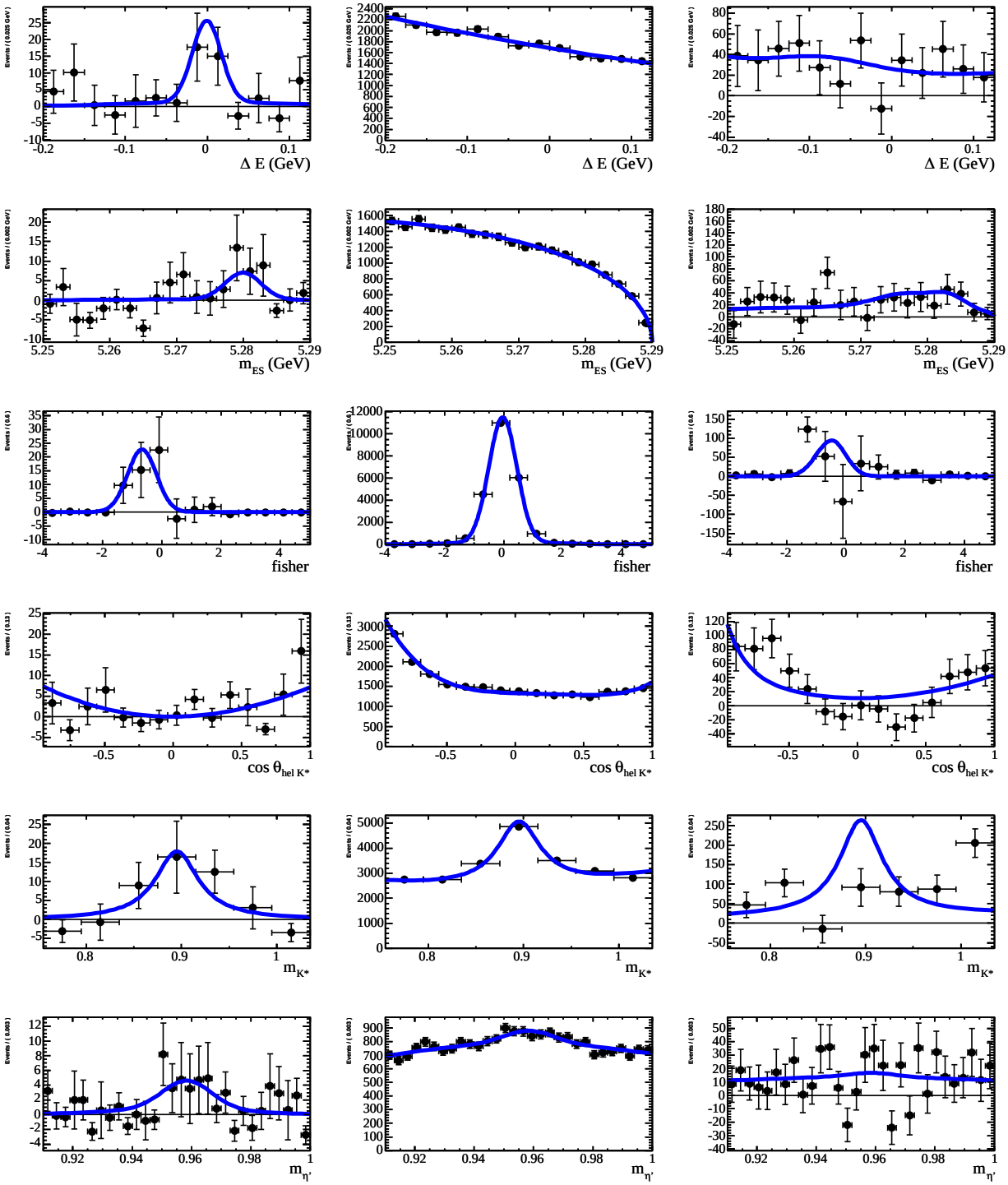

Figure B.11: sPlots for $\eta'_{\rho\gamma}K^{*0}$. The columns are signal (left), $q\overline{q}$ background (middle), and $B\overline{B}$ background (right).

B.5.3 $\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$

sPlot for	nBkg (fit)	nChls (fit)	nSig (fit)
ΔE	2086 ± 47	14 ± 10	14 ± 6
m_{ES}	2085 ± 47	15 ± 10	14 ± 6
$\mathcal{F}$	2077 ± 49	25 ± 18	12 ± 7
$\cos\theta_H$	2077 ± 49	24 ± 20	13 ± 6
m^{K^*}	2086 ± 47	16 ± 10	12 ± 5
$m^{\eta'}$	2082 ± 47	15 ± 10	17 ± 7

Table B.7: Fit yields for $q\overline{q}$ background, charmless $B\overline{B}$ background and signal observed while making sPlots for the mode $\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$, effectively removing one discriminating variable at a time from the fit.

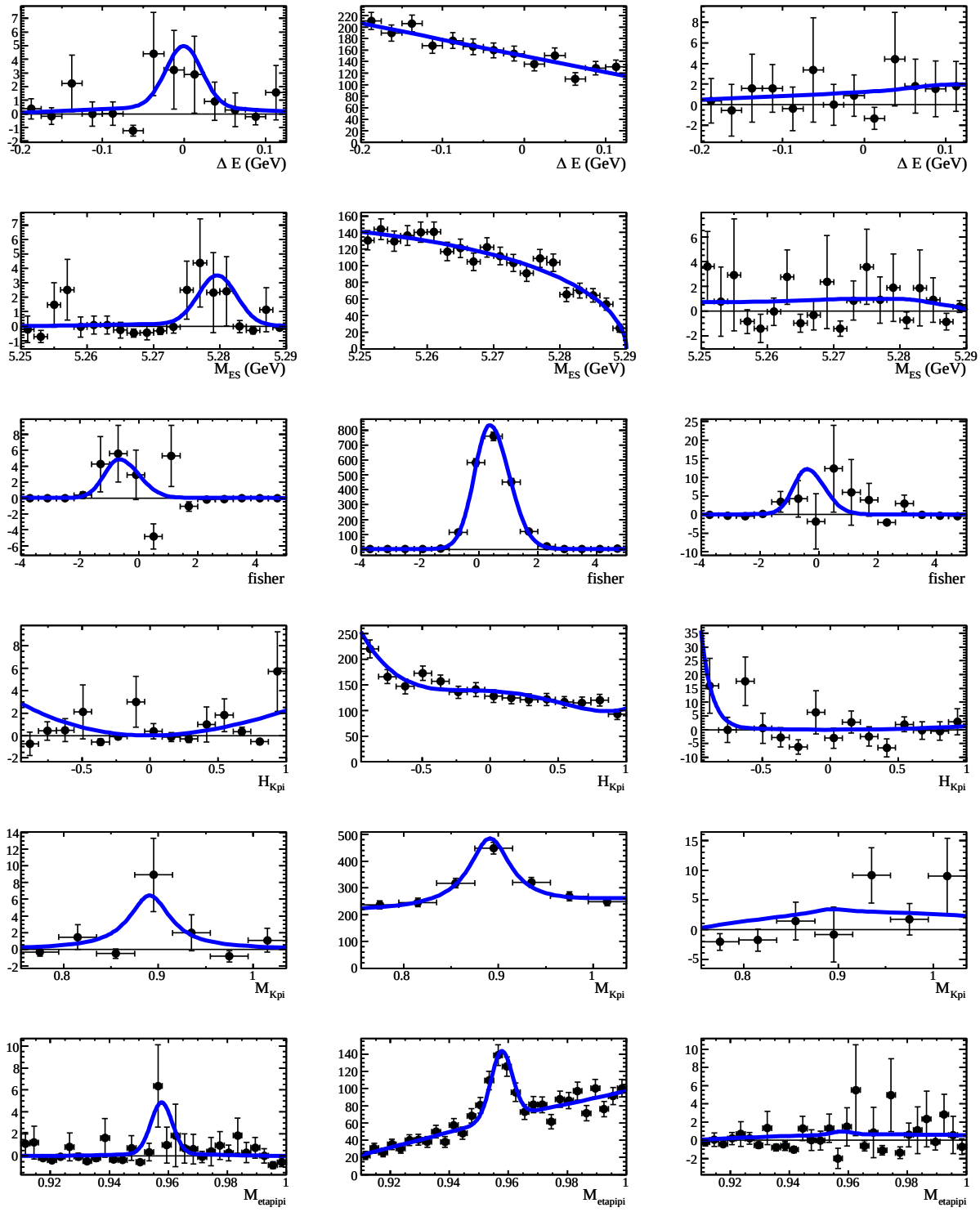

Figure B.12: sPlots for $\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$. The columns are signal (left), $q\overline{q}$ background (middle), and $B\overline{B}$ background (right).

B.5.4 $\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$

sPlot for	nBkg (fit)	nChls (fit)	nSig (fit)
ΔE	9742 ± 112	174 ± 62	46 ± 20
m_{ES}	9709 ± 114	246 ± 63	8 ± 15
$\mathcal{F}$	9826 ± 128	115 ± 84	22 ± 14
$\cos\theta_H$	9688 ± 118	256 ± 70	18 ± 12
m^{K^*}	9712 ± 114	233 ± 63	16 ± 12
$m^{\eta'}$	9728 ± 111	209 ± 57	25 ± 14

Table B.8: Fit yields for $q\overline{q}$ background, charmless $B\overline{B}$ background and signal observed while making sPlots for the mode $\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$, effectively removing one discriminating variable at a time from the fit.

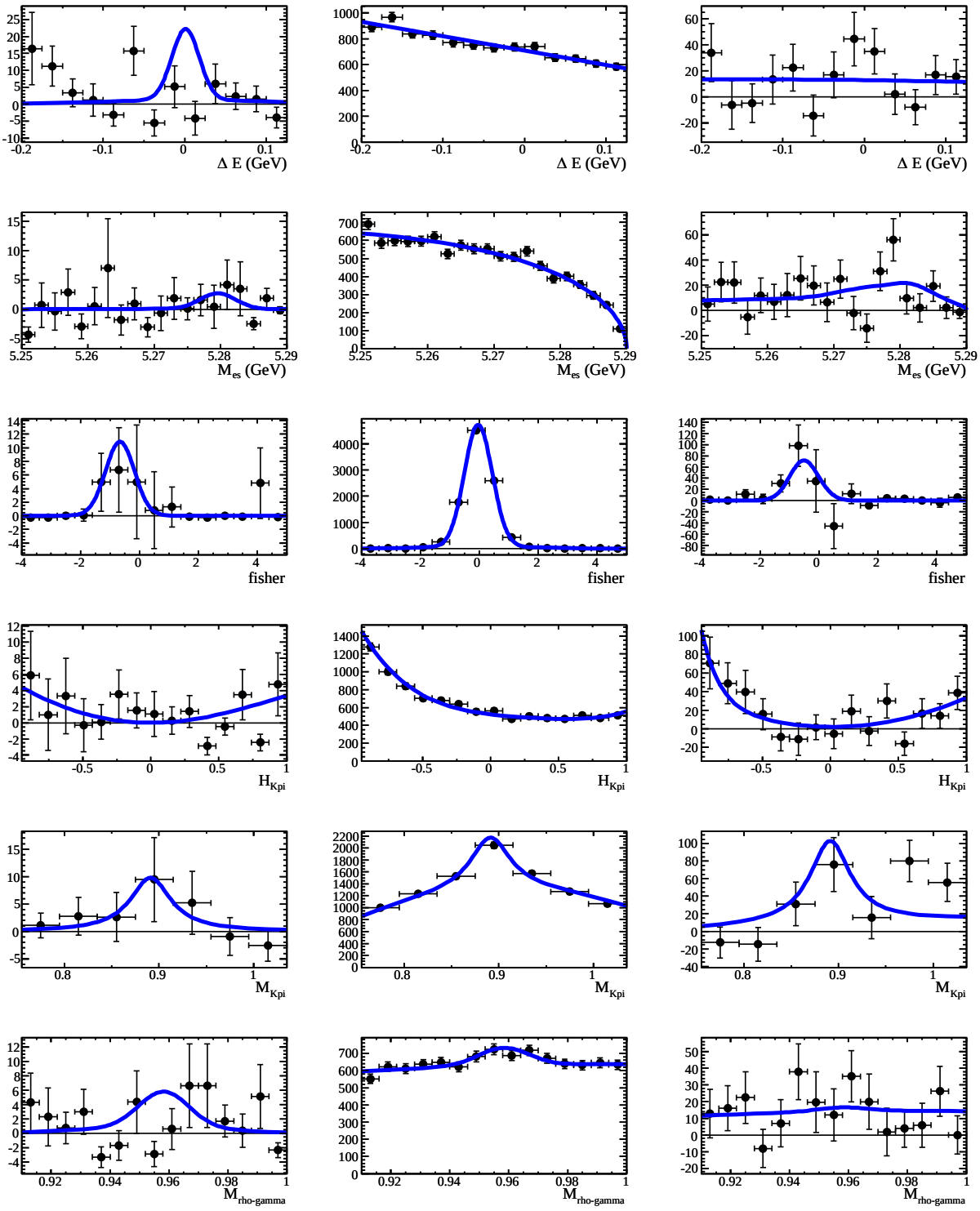

Figure B.13: sPlots for $\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$. The columns are signal (left), $q\overline{q}$ background (middle), and $B\overline{B}$ background (right).

B.6 Fit validation

B.6.1 Pure toy studies

Mode	N_{total}	N_{sig} (in)	$N_{B\bar{B}}$ (in)	N_{sig} (fit)	$\sigma(N_{sig})$ (fit)	Bias [evts]
$\eta'_{\eta\pi\pi}K^{*0}$	4837	0	22	-3.5 ± 0.2	3.9	-3.5 ± 0.2
	4837	10	22	9.5 ± 0.2	7.3	-0.5 ± 0.2
$\eta'_{\rho\gamma}K^{*0}$	23790	0	191	-2.3 ± 0.4	10.1	-2.3 ± 0.4
	23790	10	191	8.8 ± 0.4	11.3	-1.2 ± 0.4
$\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$	2114	0	11	-4.7 ± 0.3	2.9	-4.3 ± 0.2
	2114	10	11	9.8 ± 0.2	4.7	-0.2 ± 0.3
$\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$	9962	0	68	-2.0 ± 0.3	7.4	-2.0 ± 0.3
	9962	10	68	9.6 ± 0.3	8.8	-0.4 ± 0.3

Table B.9: Summary of results from pure toy MC studies for 0 and 10 signal events, and the expected number of charmless $B\overline{B}$ events included in each sample. In each case 200–1,000 toy experiments have been used. The mean N_{sig} is taken from the average over all experiments, and its uncertainty is determined as the RMS of the distribution divided by $\sqrt{\text{N Experiments}}$. The value of $\sigma(N_{sig})$ is the average uncertainty for the fitted signal yield.

B.6.2 Embedded toy studies

Mode	N_{total}	N_{sig} (in)	$N_{B\overline{B}}$ (in)	N_{sig} (fit)	$N_{B\overline{B}}$ (fit)	$\sigma(N_{sig})$ (fit)	$\sigma(N_{B\overline{B}})$ (fit)	bias [evts]
$\eta'_{\eta\pi\pi}K^{*0}$	4837	0	22	-3.7 ± 0.3	21.8 ± 0.9	3.8	19.9	-3.7 ± 0.3
	4837	10	22	11.1 ± 0.2	19.1 ± 0.9	5.7	20.2	$+1.1 \pm 0.2$
no $B\overline{B}$ fit	4837	10	22	11.6 ± 0.2	–	5.8	–	$+1.6 \pm 0.2$
nFit-bias	4837	21.3	44.5	23.0 ± 0.3	46.1 ± 0.9	7.4	22.4	$+1.7 \pm 0.3$
$\eta'_{\rho\gamma}K^{*0}$	23790	0	191	-0.0 ± 0.5	201.1 ± 3.2	10.3	115.7	-0.0 ± 0.5
	23790	10	191	13.8 ± 0.5	198.4 ± 3.5	11.7	116.4	$+3.8 \pm 0.5$
no $B\overline{B}$ fit	23790	10	191	18.1 ± 0.5	–	12.03	–	$+8.1 \pm 0.5$
nFit-bias	23790	25	409	34.5 ± 0.6	400.2 ± 5.2	14.2	122.7	$+9.5 \pm 0.6$
$\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$	2114	0	11	-4.5 ± 0.3	11	2.9	0	-4.5 ± 0.3
	2114	10	11	10.7 ± 0.2	11	4.8	0	$+0.7 \pm 0.2$
no $B\overline{B}$ fit	2114	10	11	11.1 ± 0.2	–	4.9	–	$+1.1 \pm 0.2$
nFit-bias	2114	10.8	16	11.6 ± 0.2	16.2 ± 0.5	3.8	10.7	$+0.8 \pm 0.2$
$\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$	9962	0	68	-1.4 ± 0.4	76.3 ± 2.9	7.4	70.7	-1.4 ± 0.4
	9962	10	68	12.1 ± 0.4	70.6 ± 2.8	9.1	72.3	$+2.1 \pm 0.4$
no $B\overline{B}$ fit	9962	10	68	14.4 ± 0.4	–	8.9	–	$+4.4 \pm 0.4$
nFit-bias	9962	12	231	14.9 ± 0.4	228.4 ± 2.8	9.7	70.2	$+2.9 \pm 0.4$

Table B.10: Summary of embedded toy experiments. We include an additional line called "nFit-bias". This is a set of toys which give yields as close as possible to the yields found from the final fit. We use this for the final fit bias and systematic.

B.7 Fit details- floated parameter values

B.7.1 $\eta'_{\eta\pi\pi}K^{*0}$

Correlation matrices

Table B.11: Correlation matrices for signal MC, on-peak data and $B\bar{B}$ background MC for the mode $\eta'_{\eta\pi\pi}K^{*0}$.

	$\mathcal{F}$	m_{K^*}	$\cos\theta_{\mathcal{H}}$	ΔE	m_{ES}
Signal MC:					
m_{K^*}	-0.0037				
$\cos\theta_{\mathcal{H}}$	0.0017	-0.0085			
ΔE	-0.0366	0.0127	0.0132		
m_{ES}	-0.0534	-0.0200	0.0751	0.1080	
$\mathrm{m}_{\eta'}$	0.0124	0.0070	0.0025	0.0132	-0.0155
On-Peak data:					
m_{K^*}	0.0060				
$\cos\theta_{\mathcal{H}}$	0.0607	-0.0215			
ΔE	-0.0365	0.0113	-0.0378		
m_{ES}	0.0216	0.0022	-0.0109	-0.0203	
$\mathrm{m}_{\eta'}$	0.0041	0.0141	0.0277	-0.0429	0.0270
$B\bar{B}$MC:					
m_{K^*}	0.0520				
$\cos\theta_{\mathcal{H}}$	-0.0145	-0.2053			
ΔE	-0.0488	-0.0321	-0.0599		
m_{ES}	-0.0330	-0.0726	0.1095	0.0981	
$\mathrm{m}_{\eta'}$	-0.0277	0.0206	-0.0552	0.0999	0.0158

Parameter values and pulls from pure toy studies

Table B.12: Parameter values determined from pure toy studies for the mode $\eta'_{\eta\pi\pi}K^{*0}$. Where errors are signified as zero, they are < 0.005.

Parameter	mean	sigma
deBkgP01	-1.47 ± 0.01	0.20 ± 0.00
deBkgP02	-1.06 ± 0.06	1.90 ± 0.04
fisBkgCasym	0.01 ± 0.00	0.03 ± 0.00
fisBkgCmean	0.44 ± 0.00	0.01 ± 0.00
fisBkgCrms	0.56 ± 0.00	0.01 ± 0.00
helKstarBkgP01	-0.08 ± 0.00	0.03 ± 0.00
mEpBkgfracEp	0.13 ± 0.00	0.01 ± 0.00
mEpPolyBkgP01	0.57 ± 0.00	0.02 ± 0.00
mEpPolyBkgP02	0.01 ± 0.00	0.03 ± 0.00
mKstarBkgfracKst	0.21 ± 0.00	0.02 ± 0.00
mKstarPolyBkgP01	0.05 ± 0.00	0.03 ± 0.00
mKstarPolyBkgP02	-0.07 ± 0.00	0.04 ± 0.00
mesBkgc	-29.01 ± 0.11	3.61 ± 0.08
nBkg	4805.40 ± 2.33	73.57 ± 1.67
nChmls	21.60 ± 0.68	21.36 ± 0.48
nSig	9.49 ± 0.20	6.17 ± 0.14

Table B.13: Parameter pulls determined from pure toy studies for the mode $\eta'_{\eta\pi\pi}K^{*0}$.

Pulls	mean	σ_{fit}	mean err
deBkgP01	-0.10 ± 0.03	1.02 ± 0.02	0.1971
deBkgP02	-0.10 ± 0.03	1.01 ± 0.02	1.8894
fisBkgCasym	-0.02 ± 0.03	0.98 ± 0.02	0.0257
fisBkgCmean	-0.00 ± 0.03	0.98 ± 0.02	0.0092
fisBkgCrms	-0.03 ± 0.03	1.01 ± 0.02	0.0068
helKstarBkgP01	-0.00 ± 0.03	0.98 ± 0.02	0.0297
mEpBkgfracEp	-0.02 ± 0.03	0.96 ± 0.02	0.0095
mEpPolyBkgP01	-0.06 ± 0.03	1.02 ± 0.02	0.0243
mEpPolyBkgP02	0.01 ± 0.03	0.98 ± 0.02	0.0281
mKstarBkgfracKst	-0.05 ± 0.03	1.05 ± 0.02	0.0223
mKstarPolyBkgP01	0.06 ± 0.03	1.02 ± 0.02	0.0294
mKstarPolyBkgP02	-0.05 ± 0.03	1.03 ± 0.02	0.0434
mesBkgc	0.01 ± 0.03	1.00 ± 0.02	3.6225
nBkg	-0.00 ± 0.03	1.02 ± 0.02	72.1435
nChmls	-0.10 ± 0.03	1.06 ± 0.02	20.5784
nSig	-0.30 ± 0.04	1.28 ± 0.03	5.5794

Maximum likelihood fit results

Table B.14: Maximum likelihood fit results for $\eta'_{\eta\pi\pi}K^{*0}$. The parameter GblCorr is the global correlation coefficient, defined for a particular variable as the linear combination of correlations which maximises the correlation between that variable and all others.

Floating Parameter	InitialValue	FinalValue	(+HiError,-LoError)	GblCorr.
deBkgP01	-1.4519	-1.6839	$(+0.202, -0.205)$	0.516322
deBkgP02	-0.93859	0.15101	$(+2.01, -1.94)$	0.506539
fisBkgCasym	0.0073220	0.025487	$(+0.0275, -0.0261)$	0.199849
fisBkgCmean	0.43707	0.45656	$(+0.00992, -0.00905)$	0.356562
fisBkgCrms	0.55730	0.56039	$(+0.00676, -0.00715)$	0.251214
helKstarBkgP01	-0.076223	-0.079754	$(+0.0315, -0.0301)$	0.152609
mEpBkgfracEp	0.12732	0.12603	$(+0.00960, -0.00963)$	0.406744
mEpPolyBkgP01	0.57585	0.60926	$(+0.0244, -0.0244)$	0.173728
mEpPolyBkgP02	0.012929	0.026346	$(+0.0277, -0.0279)$	0.427669
mKstarBkgfracKst	0.21047	0.21976	$(+0.0226, -0.0227)$	0.719302
mKstarPolyBkgP01	0.044941	0.066611	$(+0.0300, -0.0303)$	0.083273
mKstarPolyBkgP02	-0.065263	-0.039357	$(+0.0451, -0.0445)$	0.719852
mesBkgc	-29.041	-21.080	$(+3.75, -3.65)$	0.088297
nBkg	4,805.0	4,771.7	$(+70.6, -76.0)$	0.284468
nChmls	22.000	42.602	$(+27.7, -19.4)$	0.506259
nSig	10.000	22.392	$(+8.00, -6.55)$	0.177152

B.7.2 $\eta'_{\rho\gamma}K^{*0}$

Correlation matrices

Table B.15: Correlation matrices for signal MC, on-peak data and $B\bar{B}$ background MC for the mode $\eta'_{\rho\gamma}K^{*0}$.

	$\mathcal{F}$	m_{K^*}	$\cos\theta_{\mathcal{H}}$	ΔE	m_{ES}
Signal MC:					
m_{K^*}	-0.0023				
$\cos\theta_{\mathcal{H}}$	-0.0006	-0.0166			
ΔE	-0.0361	0.0377	0.0042		
m_{ES}	-0.0507	-0.0258	0.0733	0.1034	
$\mathrm{m}_{\eta'}$	0.0078	-0.0010	-0.0114	0.0230	0.0101
On-Peak data:					
m_{K^*}	-0.0116				
$\cos\theta_{\mathcal{H}}$	0.0373	-0.0383			
ΔE	-0.0112	-0.0120	0.0093		
m_{ES}	-0.0120	0.0037	0.0061	-0.0085	
$\mathrm{m}_{\eta'}$	-0.0045	0.0113	0.0036	0.0017	0.0004
$B\bar{B}$MC:					
m_{K^*}	0.0006				
$\cos\theta_{\mathcal{H}}$	-0.0369	-0.0384			
ΔE	0.0076	-0.0157	-0.0782		
m_{ES}	-0.0423	-0.0197	0.1022	0.0130	
$\mathrm{m}_{\eta'}$	-0.0049	0.0091	-0.0200	-0.0577	0.0146

Parameter values and pulls from pure toy studies

Table B.16: Parameter values determined from pure toy studies for the mode $\eta'_{\rho\gamma}K^{*0}$. Where errors are signified as zero, they are < 0.005.

Parameter	mean	sigma
deBkgP01	-1.42 ± 0.00	0.09 ± 0.00
deBkgP02	1.00 ± 0.03	0.85 ± 0.02
fisBkgCasym	-0.00 ± 0.00	0.01 ± 0.00
fisBkgCmean	-0.04 ± 0.00	0.00 ± 0.00
fisBkgCrms	0.47 ± 0.00	0.00 ± 0.00
helKstarBkgP01	-0.09 ± 0.00	0.01 ± 0.00
mEpBkgfracEp	0.03 ± 0.00	0.01 ± 0.00
mEpPolyBkgP01	0.02 ± 0.00	0.01 ± 0.00
mEpPolyBkgP02	-0.06 ± 0.00	0.02 ± 0.00
mKstarBkgfracKst	0.23 ± 0.00	0.01 ± 0.00
mKstarPolyBkgP01	0.06 ± 0.00	0.01 ± 0.00
mKstarPolyBkgP02	0.05 ± 0.00	0.02 ± 0.00
mesBkgc	-24.36 ± 0.05	1.68 ± 0.04
nBkg	23578.55 ± 5.37	169.60 ± 3.82
nChmls	195.14 ± 2.67	84.50 ± 1.92
nSig	8.78 ± 0.37	11.85 ± 0.27

Table B.17: Parameter pulls determined from pure toy studies for the mode $\eta'_{\rho\gamma}K^{*0}$.

Pulls	mean	σ_{fit}	mean err
deBkgP01	-0.04 ± 0.03	1.00 ± 0.02	0.0908
deBkgP02	-0.04 ± 0.03	0.96 ± 0.02	0.8878
fisBkgCasym	0.07 ± 0.03	0.98 ± 0.02	0.0118
fisBkgCmean	0.06 ± 0.03	0.97 ± 0.02	0.0040
fisBkgCrms	-0.02 ± 0.03	0.97 ± 0.02	0.0027
helKstarBkgP01	-0.01 ± 0.03	1.01 ± 0.02	0.0139
mEpBkgfracEp	-0.00 ± 0.03	0.97 ± 0.02	0.0096
mEpPolyBkgP01	0.04 ± 0.03	1.03 ± 0.02	0.0117
mEpPolyBkgP02	-0.04 ± 0.03	0.96 ± 0.02	0.0174
mKstarBkgfracKst	-0.01 ± 0.03	0.99 ± 0.02	0.0103
mKstarPolyBkgP01	-0.03 ± 0.03	1.00 ± 0.02	0.0134
mKstarPolyBkgP02	-0.01 ± 0.03	1.01 ± 0.02	0.0200
mesBkgc	-0.02 ± 0.03	0.96 ± 0.02	1.7484
nBkg	-0.06 ± 0.03	0.89 ± 0.02	190.3302
nChmls	0.03 ± 0.02	0.73 ± 0.02	115.9531
nSig	-1.38 ± 0.27	1.91 ± 0.18	11.3283

Maximum likelihood fit results

Table B.18: Maximum likelihood fit results for $\eta'_{\rho\gamma}K^{*0}$. The parameter GblCorr is the global correlation coefficient, defined for a particular variable as the linear combination of correlations which maximises the correlation between that variable and all others.

Floating Parameter	InitialValue	FinalValue	(+HiError,-LoError)	GblCorr.
deBkgP01	-1.4209	-1.4827	$(+0.0916, -0.0909)$	0.513350
deBkgP02	1.0252	1.1790	$(+0.901, -0.891)$	0.512796
fisBkgCasym	-0.0032941	0.0042985	$(+0.0119, -0.0119)$	0.316533
fisBkgCmean	-0.040863	-0.035961	$(+0.00413, -0.00413)$	0.569510
fisBkgCrms	0.47245	0.47261	$(+0.00283, -0.00282)$	0.379244
helKstarBkgP01	-0.088202	-0.084880	$(+0.0141, -0.0141)$	0.019270
mEpBkgfracEp	0.034877	0.041146	$(+0.00961, -0.00939)$	0.745921
mEpPolyBkgP01	0.015124	0.010806	$(+0.0118, -0.0119)$	0.114209
mEpPolyBkgP02	-0.056110	-0.047831	$(+0.0174, -0.0170)$	0.744124
mKstarBkgfracKst	0.23132	0.20669	$(+0.0103, -0.0102)$	0.753664
mKstarPolyBkgP01	0.064426	0.065444	$(+0.0132, -0.0132)$	0.059231
mKstarPolyBkgP02	0.048043	0.053246	$(+0.0197, -0.0195)$	0.739463
mesBkgc	-24.327	-20.993	$(+1.76, -1.77)$	0.340854
nBkg	$23,595$	$23,360$	$(+190, -189)$	0.571397
nChmls	185.00	394.80	$(+116, -114)$	0.763000
nSig	10.000	34.999	$(+14.3, -12.7)$	0.239548

B.7.3 $\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$

Correlation matrices

Table B.19: Correlation matrices for signal MC, on-peak data and $B\bar{B}$ background MC for the mode $\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$.

	$\mathcal{F}$	m_{K^*}	$\cos\theta_{\mathcal{H}}$	ΔE	m_{ES}
Signal MC:					
m_{K^*}	-0.0012				
$\cos\theta_{\mathcal{H}}$	-0.0122	-0.0179			
ΔE	-0.0335	0.0327	-0.0119		
m_{ES}	-0.0509	-0.0109	0.0809	0.1186	
$m_{\eta'}$	0.0172	-0.0029	-0.0041	-0.0110	-0.0326
On-Peak data:					
m_{K^*}	-0.0164				
$\cos\theta_{\mathcal{H}}$	-0.0043	-0.0374			
ΔE	-0.0425	0.0424	-0.0107		
m_{ES}	-0.0185	0.0484	0.0110	-0.0421	
$m_{\eta'}$	0.0192	0.0129	-0.0353	-0.0270	-0.0001
$B\bar{B}$MC:					
m_{K^*}	0.0530				
$\cos\theta_{\mathcal{H}}$	-0.0050	-0.0954			
ΔE	-0.1162	-0.0213	-0.1086		
m_{ES}	-0.0116	0.0041	0.1011	0.0842	
$m_{\eta'}$	-0.0213	0.0181	0.0164	0.0466	0.0297

Parameter values and pulls from pure toy studies

Table B.20: Parameter values determined from pure toy studies for the mode $\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$. Where errors are signified as zero, they are < 0.005.

Parameter	mean	sigma
deBkgP01	-1.72 ± 0.01	0.27 ± 0.01
fisBkgCasym	0.07 ± 0.00	0.04 ± 0.00
fisBkgCmean	0.47 ± 0.00	0.01 ± 0.00
fisBkgCrms	0.58 ± 0.00	0.01 ± 0.00
mEpBkgfracEp	0.13 ± 0.00	0.01 ± 0.00
mEpPolyBkgP01	0.60 ± 0.00	0.03 ± 0.00
mKstarBkgfracKst	0.20 ± 0.00	0.02 ± 0.00
mKstarPolyBkgP01	0.02 ± 0.00	0.04 ± 0.00
mesBkgc	-22.15 ± 0.18	5.73 ± 0.13
nBkg	2090.42 ± 1.47	46.42 ± 1.06
nSig	9.76 ± 0.16	5.14 ± 0.12

Table B.21: Parameter pulls determined from pure toy studies for the mode $\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$.

Pulls	mean	σ_{fit}	mean err
deBkgP01	-0.02 ± 0.03	1.02 ± 0.02	0.2595
fisBkgCasym	-0.00 ± 0.03	1.02 ± 0.02	0.0361
fisBkgCmean	0.02 ± 0.03	0.99 ± 0.02	0.0133
fisBkgCrms	-0.01 ± 0.03	0.99 ± 0.02	0.0098
mEpBkgfracEp	-0.03 ± 0.03	0.99 ± 0.02	0.0133
mEpPolyBkgP01	-0.01 ± 0.03	0.96 ± 0.02	0.0360
mKstarBkgfracKst	-0.00 ± 0.03	1.02 ± 0.02	0.0224
mKstarPolyBkgP01	0.05 ± 0.03	0.98 ± 0.02	0.0441
mesBkgc	-0.02 ± 0.03	1.05 ± 0.02	5.4828
nBkg	-0.07 ± 0.03	1.01 ± 0.02	45.9599
nSig	-0.26 ± 0.04	1.22 ± 0.03	4.6960

Maximum likelihood fit results

Table B.22: Maximum likelihood fit results for $\eta'_{\eta\pi\pi}K^{*+}_{K^0\pi^+}$. The parameter GblCorr is the global correlation coefficient, defined for a particular variable as the linear combination of correlations which maximises the correlation between that variable and all others.

Floating Parameter	InitialValue	FinalValue	(+HiError,-LoError)	GblCorr.
deBkgP01	-1.7114	-1.8934	$(+0.259, -0.262)$	0.098722
fisBkgCasym	0.069488	0.11679	$(+0.0389, -0.0387)$	0.126653
fisBkgCmean	0.47234	0.47608	$(+0.0138, -0.0137)$	0.247447
fisBkgCrms	0.57974	0.58565	$(+0.0102, -0.00993)$	0.200950
mEpBkgfracEp	0.12740	0.13696	$(+0.0136, -0.0134)$	0.076251
mEpPolyBkgP01	0.60174	0.61851	$(+0.0357, -0.0366)$	0.022885
mKstarBkgfracKst	0.20184	0.22472	$(+0.0227, -0.0226)$	0.065414
mKstarPolyBkgP01	0.019792	0.085641	$(+0.0451, -0.0452)$	0.081370
mesBkgc	-22.087	-16.345	$(+5.55, -5.53)$	0.072352
nBkg	$2,094.0$	$2,086.3$	$(+47.1, -46.5)$	0.190881
nChmls	16.000	16.434	$(+11.2, -9.34)$	0.337113
nSig	11.000	11.209	$(+5.68, -4.50)$	0.153372

B.7.4 $\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$

Correlation matrices

Table B.23: Correlation matrices for signal MC, on-peak data and $B\bar{B}$ background MC for the mode $\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$.

	$\mathcal{F}$	m_{K^*}	$\cos\theta_{\mathcal{H}}$	ΔE	m_{ES}
Signal MC:					
m_{K^*}	-0.0003				
$\cos\theta_{\mathcal{H}}$	-0.0252	-0.0266			
ΔE	-0.0276	0.0402	-0.0050		
m_{ES}	-0.0623	-0.0171	0.0855	0.1049	
$m_{\eta'}$	0.0172	0.0056	-0.0022	0.0472	0.0058
On-Peak data:					
m_{K^*}	-0.0113				
$\cos\theta_{\mathcal{H}}$	0.0352	-0.0617			
ΔE	0.0019	0.0083	0.0211		
m_{ES}	-0.0058	-0.0117	0.0124	-0.0038	
$m_{\eta'}$	-0.0011	-0.0035	-0.0000	-0.0094	0.0090
$B\bar{B}$MC:					
m_{K^*}	0.0013				
$\cos\theta_{\mathcal{H}}$	-0.0654	-0.1260			
ΔE	-0.0181	0.0472	-0.1400		
m_{ES}	-0.0851	-0.0038	0.1227	0.0380	
$m_{\eta'}$	-0.0295	-0.0334	0.0061	-0.0304	0.0101

Parameter values and pulls from pure toy studies

Table B.24: Parameter values determined from pure toy studies for the mode $\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$. Where errors are signified as zero, they are < 0.005.

Parameter	mean	sigma
deBkgP01	-1.48 ± 0.00	0.12 ± 0.00
fisBkgCasym	-0.00 ± 0.00	0.02 ± 0.00
fisBkgCmean	-0.04 ± 0.00	0.01 ± 0.00
fisBkgCrms	0.47 ± 0.00	0.00 ± 0.00
mEpBkgfracEp	0.03 ± 0.00	0.01 ± 0.00
mEpPolyBkgP01	0.03 ± 0.00	0.02 ± 0.00
mKstarBkgfracKst	0.17 ± 0.00	0.01 ± 0.00
mKstarPolyBkgP01	0.11 ± 0.00	0.02 ± 0.00
mesBkgc	-25.32 ± 0.08	2.55 ± 0.06
nBkg	9879.00 ± 3.64	115.04 ± 2.60
nChmls	72.47 ± 1.93	61.13 ± 1.37
nSig	9.64 ± 0.31	9.64 ± 0.22

Table B.25: Parameter pulls determined from pure toy studies for the mode $\eta'_{\rho\gamma} K^{*+}_{K^0\pi^+}$.

Pulls	mean	σ_{fit}	mean err
deBkgP01	0.02 ± 0.03	0.99 ± 0.02	0.1182
fisBkgCasym	0.04 ± 0.03	1.01 ± 0.02	0.0185
fisBkgCmean	0.06 ± 0.03	0.98 ± 0.02	0.0060
fisBkgCrms	-0.03 ± 0.03	1.01 ± 0.02	0.0042
mEpBkgfracEp	0.00 ± 0.03	1.00 ± 0.02	0.0097
mEpPolyBkgP01	0.07 ± 0.03	1.00 ± 0.02	0.0178
mKstarBkgfracKst	-0.02 ± 0.03	1.01 ± 0.02	0.0115
mKstarPolyBkgP01	-0.02 ± 0.03	1.02 ± 0.02	0.0198
mesBkgc	-0.01 ± 0.03	0.97 ± 0.02	2.6216
nBkg	-0.05 ± 0.03	0.98 ± 0.02	117.5253
nChmls	0.05 ± 0.03	0.94 ± 0.02	65.2186
nSig	-0.19 ± 0.04	1.19 ± 0.03	8.7510

Maximum likelihood fit results

Table B.26: Maximum likelihood fit results for $\eta'_{\rho\gamma}K^{*+}_{K^0\pi^+}$. The parameter GblCorr is the global correlation coefficient, defined for a particular variable as the linear combination of correlations which maximises the correlation between that variable and all others.

Floating Parameter	InitialValue	FinalValue	(+HiError,-LoError)	GblCorr.
deBkgP01	-1.4791	-1.5745	$(+0.123, -0.126)$	0.047942
fisBkgCasym	-0.0018072	0.016880	$(+0.0202, -0.0179)$	0.303785
fisBkgCmean	-0.037263	-0.028973	$(+0.00743, -0.00571)$	0.536991
fisBkgCrms	0.47505	0.47407	$(+0.00421, -0.00495)$	0.328838
mEpBkgfracEp	0.030728	0.053176	$(+0.0103, -0.0102)$	0.092261
mEpPolyBkgP01	0.021710	0.034063	$(+0.0190, -0.0192)$	0.068608
mKstarBkgfracKst	0.17016	0.16747	$(+0.0112, -0.0129)$	0.291257
mKstarPolyBkgP01	0.11442	0.079146	$(+0.0205, -0.0214)$	0.096290
mesBkgc	-25.307	-20.504	$(+2.96, -2.58)$	0.276695
nBkg	$12,000$	$9,736.5$	$(+107, -141)$	0.544228
nChmls	21.000	209.08	$(+93.6, -54.1)$	0.731850
nSig	9.0000	15.668	$(+11.0, -10.7)$	0.294562

Bibliography

[1] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental Test Of Parity Conservation In Beta Decay. *Phys. Rev.*, 105:1413–1414, 1957.

[2] T. D. Lee and C. N. Yang. Question of Parity Conservation in Weak Interactions. *Phys. Rev.*, 104:254–258, Oct 1956.

[3] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2π Decay of the K_2^0 Meson. *Phys. Rev. Lett.*, 13:138–140, Jul 1964.

[4] A. Sakharov. Violation of CP Symmetry, C-Asymmetry and Baryon Asymmetry of the Universe. *Pisma Zh. Eksp. Teor. Fiz.*, 5:32–35, Jan 1967.

[5] B. Aubert et al. Observation of CP Violation in the B^0 Meson System. *Phys. Rev. Lett.*, 87:091801, Aug 2001.

[6] K. Abe et al. Observation of Large CP Violation in the Neutral B Meson System. *Phys. Rev. Lett.*, 87:091802, Aug 2001.

[7] B. Aubert et al. Direct CP Violating Asymmetry in $B^0 \to K^+\pi^-$ Decays. *Phys. Rev. Lett.*, 93:131801, Sep 2004.

[8] Y. Chao et al. Evidence for Direct CP Violation in $B^0 \to K^+\pi^-$ Decays. *Phys. Rev. Lett.*, 93:191802, 2004.

[9] N. Cabibbo. Unitary Symmetry and Leptonic Decays. *Phys. Rev. Lett.*, 10:531–533, Jun 1963.

[10] S.L. Glashow, J. Iliopoulos, and L. Maiani. Weak Interactions with Lepton-Hadron Symmetry. *Phys. Rev.*, D2:1285–1292, 1970.

[11] S.L. Glashow. Partial Symmetries of Weak Interactions. *Nucl.Phys.*, 22:579–588, 1961.

[12] M. Kobayashi and T. Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[13] S. Eidelman et al. Review of particle physics. Particle Data Group. *Phys.Lett.*, B592:1, 2004.

[14] F.J. Gilman and R. Kauffman. $\eta - \eta'$ Mixing Angle. *Phys. Rev. D*, 36:2761–2767, Nov 1987.

[15] A. Datta, H.J. Lipkin, and P.J. O'Donnell. Nonstandard η - η' Mixing and the Nonleptonic B and Λ_b Decays to η and η'. *Phys.Lett.*, B544:145–153, 2002.

[16] B. Aubert et al. B meson decays to $\eta^{(\prime)}K^*$, $\eta^{(\prime)}\rho$, $\eta^{(\prime)}\pi^0$, $\omega\pi^0$, and $\phi\pi^0$. *Phys. Rev. D*, 70:032006, Aug 2004.

[17] K. Anikeev et al. Averages of $b-$hadron properties as of Winter 2005. *Heavy Flavor Averaging Group (HFAG)*, 2005.

[18] C.W. Chiang, M. Gronau, Z. Luo, J.L. Rosner, and D.A. Suprun. Charmless $B \to VP$ decays using flavor SU(3) symmetry. *Phys.Rev.*, D69:034001, 2004.

[19] M. Beneke and M. Neubert. QCD factorization for $B \to PP$ and $B \to PV$ decays. *Nucl.Phys.*, B675:333–415, 2003.

[20] Hiroaki Aihara. Charmless B decays involving η, η' and vector mesons. *eConf*, C030603:MAR12, 2003.

[21] J. Iizuka. Systematics and phenomenology of meson family. *Prog.Theor.Phys.Suppl.*, 37:21–34, 1966.

[22] M. Beneke and M. Neubert. Flavor singlet B decay amplitudes in QCD factorization. *Nucl.Phys.*, B651:225–248, 2003.

[23] C.W. Chiang and J.L. Rosner. Updated analysis of some two-body charmless B decays. *Phys.Rev.*, D65:074035, 2002.

[24] C.W. Chiang, M. Gronau, and J.L. Rosner. Two body charmless B decays involving η and η'. *Phys.Rev.*, D68:074012, 2003.

[25] Y. Grossman, Z. Ligeti, Y. Nir, and H. Quinn. SU(3) relations and the CP asymmetries in B decays to $\eta' K_s$, ϕK_s and $K^+K^-K_s$. *Phys.Rev.*, D68:015004, 2003.

[26] M. Beneke, G. Buchalla, M. Neubert, and C.T. Sachrajda. QCD factorization for $B \to \pi\pi$ decays: Strong phases and CP violation in the heavy quark limit. *Phys.Rev.Lett.*, 83:1914–1917, 1999.

[27] M. Beneke, G. Buchalla, M. Neubert, and C.T. Sachrajda. QCD factorization for exclusive, nonleptonic B meson decays: General arguments and the case of heavy light final states. *Nucl.Phys.*, B591:313–418, 2000.

[28] X. Liu, H.S. Wang, Z.J. Xiao, L. Guo, and C.D. Lu. Branching ratio and CP asymmetry of $B \to \rho\eta'$ decays in the perturbative QCD approach. *Phys.Rev.*, D73:074002, 2006.

[29] C.W. Bauer, S. Fleming, and M.E. Luke. Summing Sudakov logarithms in $B \to X_{(s\gamma)}$ in effective field theory. *Phys.Rev.*, D63:014006, 2000.

[30] A.R. Williamson and J. Zupan. Two body B decays with isosinglet final states in SCET. *Phys.Rev.*, D74:014003, 2006.

[31] PEP-II: An Asymmetric B Factory. Conceptual Design Report. June 1993. 1994.

[32] B. Aubert et al. The *BABAR* detector. *Nucl.Instrum.Meth.*, A479:1–116, 2002.

[33] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl.Instrum.Meth.*, A506:250–303, 2003.

[34] R.A. Fisher. The use of multiple measurements in taxonomic problems. *Annals Eugen.*, 7:179–188, 1936.

[35] H. Albrecht et al. Reconstruction of B Mesons. *Phys.Lett.*, B185:218, 1987.

[36] B. Aubert et al. Observation of B^0 Meson Decay to $a^+_{1(1260)}\pi^+$. *Phys.Rev.Lett.*, 97:051802, 2006.

[37] B. Aubert et al. Measurements of branching fractions, polarizations, and direct CP-violation asymmetries in $B \to \rho K^*$ and $B \to f_{0(980)}K^*$ decays. *Phys.Rev.Lett.*, 97:201801, 2006.

[38] M. Pivk and F.R. Le Diberder. SPlot: A Statistical tool to unfold data distributions. *Nucl.Instrum.Meth.*, A555:356–369, 2005.

[39] B. Aubert et al. Observation of $B \to \eta' K^*$ and evidence for $B^+ \to \eta' \rho^+$. *Phys.Rev.Lett.*, 98:051802, 2007.

[40] J. Schumann et al. Search for B decays into $\eta'\rho$, $\eta' K^*$, $\eta'\phi$, $\eta'\omega$ and $\eta'\eta'$ at Belle. *Phys.Rev.*, D75:092002, 2007.

[41] P. del Amo Sanchez et al. B-meson decays to $\eta'\rho$, $\eta' f_0$, and $\eta' K^*$. *Phys.Rev.*, D82:011502, 2010.

[42] C.W. Chiang and Y.F. Zhou. Flavor symmetry analysis of charmless $B \to VP$ decays. *JHEP*, 0903:055, 2009.